

Playing a dangerous game

The current US administration is edging towards a policy of developing new types of nuclear weapons. In today's uncertain world, the last thing we need is a renewed arms race among the world's nuclear powers.

For the most hawkish of hawks, there's no nuke so good as a shiny, bells-and-whistles new nuke. And if you look at the fine print of the latest budget request for the US Department of Energy, it's clear that these birds of prey have been busy of late.

Any proposal to build new nuclear weapons is highly controversial — and would require the express permission of the US Congress. Yet in its projections for future years' spending, the 2005 energy department budget request is trying to sow the seeds of such developments. These projections call for "subsystem tests and a full system test" of a nuclear bomb proposed to vaporize deeply buried bunkers and other 'hardened' targets. Further on, one can find plans to rehabilitate old nuclear testing ranges in Nevada and to build a facility to mass-manufacture plutonium triggers for detonating hydrogen bombs.

This activity is geared towards what the budget describes as the "stockpile of the future" — which, if a paper authored by four nuclear-weapons scientists at the Los Alamos National Laboratory in New Mexico is anything to go by, could feature an arsenal of mini-nuclear weapons (B. L. Fearey *et al. Comp. Strategy* **22**, 305–324; 2003). The authors argue that such weapons would deter rogue states from developing subterranean stores of chemical and biological weapons. But in describing the bombs as causing "reduced collateral damage", the paper raises the spectre of 'usable' nukes — a concept that causes the colour to drain from the cheeks of arms-control advocates.

Senior officials deny that there are any concrete plans to develop new weapons — merely "paper studies" to assess future options (see page 455). But seasoned observers see the current budget request

as an attempt by the hawks within the Bush administration to pursue their new nukes agenda. These officials are backed by a minority of weapons scientists for whom life has lost meaning since their cold-war heyday, when they were given huge amounts of money to pursue the nuclear arms race with the Soviet Union.

This agenda is already damaging US national security. Russian politicians have responded with statements about developing next-generation strategic weapons. And if the United States and Russia start pursuing mini-nuclear weapons, it doesn't take much imagination to see that other countries, perhaps India and Pakistan, may follow suit. In an era in which the primary security threat seems to be from terrorism, the proliferation of portable nuclear weapons can't be a good idea.

Against this security background, America's nuclear-weapons scientists don't need to design new bombs to get back on the front lines of national defence. Many of them, for example, are members of top-secret nuclear search teams that watched over the New Year's celebrations in New York, Washington and other US cities this past winter. Others are working with the Department of Homeland Security to build detectors for ports and border crossings that can sniff out the faintest hint of radiation. And those working on stockpile-stewardship programmes continue to support global stability.

These noble efforts will be undermined by the plans of a nostalgic few. Physicists who care about arms control, and officials in the Bush administration who don't share the hawks' enthusiasm for new weapons, should unite to nip these nuclear ambitions in the bud before they make the world even more dangerous than it already is. ■

A fair deal for all

Scientists in poorer countries have to pay over the odds for equipment and reagents. They deserve a helping hand.

When your government can't afford to match the research grants paid to rich labs in North America, Europe and Japan, it must be especially galling to have to pay more than your wealthy competitors for standard lab equipment and materials.

The extent of the problem is revealed this week by a *Nature* survey of researchers in Germany, Poland, the United States and Brazil (see page 453). On both sides of the Atlantic, the larger and more competitive market of the established scientific powers ensures that prices are driven down. But elsewhere, the poor get poorer.

What can researchers do to combat these harsh economics of scale? Sometimes it helps to negotiate just a little harder. One Polish group leader told *Nature* that he got a 35% discount on centrifuges after confronting his local retailer with cheaper prices across the nearby German border. But Polish scientists have an ace up their sleeve: their country's imminent membership of the European Union means that they will soon be able to complain to officials in Brussels if they feel they are not benefiting from the continent's common market.

For scientists in other poorer countries, the best answer would be to negotiate transnational agreements to ensure that every research group similarly benefits from being part of a large, international market. Unfortunately, scientists and the bodies representing their

interests can do little to affect the politics of international trade. But that doesn't mean we must accept the iniquitous status quo. The Stockholm-based International Foundation for Science, a non-governmental organization that aims to help scientists in developing countries, provides one example: its grants, which totalled US\$2.5 million in 2003, included a \$580,000 purchasing service for scientific instruments and reagents. And the Pew Charitable Trusts, based in Philadelphia, have recently set up an initiative to make second-hand equipment available to their Latin American research fellows.

These are encouraging signs, but more and larger efforts are needed. In particular, there is a case for the United Nations Educational, Scientific and Cultural Organization (UNESCO) to get more deeply involved. UNESCO already has an agreement with some US non-profit organizations to promote the redistribution of used instruments, at low prices, to labs in poorer countries. It could expand these schemes, and consider providing a purchasing service for new equipment.

Finally, the affected scientists' national governments should remove customs barriers that can severely hamper the import of scientific instruments. Our survey found cases where it took researchers more than a year to get their hands on their equipment. When you've already paid over the odds, such delays are doubly frustrating. ■

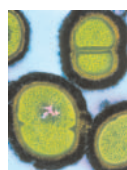




Trial run
Biohazard test kits
get put through
their paces
p454



Guiding missiles
US government
gets fired up over
nuclear weapons
p455



Shining example
Technique lights
up antibiotic-
resistant bacteria
p457



Deadly diet
Lichen linked to
death of Rocky
Mountain elk
p458

High prices of supplies drain cash from poorer nations' labs

Eva Schillinger, Munich

While researchers in the West bemoan budget cuts and funding constraints, their colleagues outside this wealthy enclave are struggling to pay frequently inflated prices for lab equipment and materials.

Scientists in these poorer countries have to pay up to 70% more than their wealthier colleagues for identical supplies, a survey by *Nature* has revealed.

In Poland, for example, researchers pay two-thirds more for an EPS 601 power supply for gel separation manufactured by Amersham Biosciences than their neighbours over the border in Germany (see Table). And in Brazil, a kilogram of yeast extract produced by Sigma-Aldrich of St Louis, Missouri, costs at least US\$340 — 70% more than the price listed in the United States.

For a bench-top 5415 D centrifuge made by Eppendorf in Hamburg, Germany — a typical piece of equipment for life-sciences labs — Brazilians have to pay US\$3,110, almost 60% more than their US colleagues.

Suppliers point out that they do not have control of prices in individual countries, which are set by local distributors. The discrepancies are explained by market conditions, the suppliers note — in particular, the higher costs of doing business in smaller, less-established markets.

"Price differences can exist for certain products due to many factors," says Ron Hutton, treasurer of equipment supplier Bio-Rad Laboratories of Hercules, California, "including differences in support provided, quantities, shipping costs, currency values and local competitive pressures." Jörn Peplow, a spokesman for Eppendorf, says that pricing "depends on the distribution channel". Mike Hogan, chief financial officer of Sigma-Aldrich, adds: "It's the cost of doing business."

But disadvantaged researchers take a less sanguine view. "It is a big problem, everything is so much more expensive in Poland," says Jan Potempa, a microbiologist with positions at both Jagiellonian University in Kraków and the University of Georgia, Athens, in the United States. "We just need two-thirds more money."

Nature's survey was triggered by private

Uneven spread: the cost of lab equipment and materials (US\$)

	 Germany	 Poland	 United States	 Brazil
Eppendorf Centrifuge 5415 D	\$1,780	\$1,960	\$1,950	\$3,110
Bio-Rad Laboratories Mini-PROTEAN 3 cell kit	\$1,210	\$990	\$510	\$670
Amersham Biosciences EPS 601 power supply	\$1,030	\$1,720	\$960	\$1,420
Amersham Biosciences Ultraspec 3300 pro spectrometer	\$10,420	\$16,780	\$12,190	\$16,000
Promega Taq DNA polymerase (100 units)	\$54	\$72	\$25	\$39
Promega Sodium chloride (molecular grade, 1 kg)	\$64	\$71	\$38	\$35
Sigma-Aldrich Yeast extract (1 kg)	\$360	\$480	\$202	\$340

complaints about the pricing structure from researchers in Brazil and elsewhere. Four countries were surveyed: Poland, a major market in eastern Europe, was compared with its neighbour Germany; and Brazil, the largest Latin American market was matched against the United States. An individual scientist in each of the countries was asked to report on prices for 18 different commodities and instruments offered by local distributors for eight leading suppliers.

The survey revealed a few exceptions to the trend that scientists in poorer countries pay more. For example, a kilogram of sodium chloride from Promega in Madison, Wisconsin, cost slightly less in Brazil than in the United States. But overall, prices were higher in Poland than in Germany for 16 out of 18 products; and 11 out of the 12 products for which prices were obtained in the Americas were cheaper in the United States than in Brazil.

According to Maciej Nalecz, a molecular biologist who used to run a lab for the Nencki Institute of Experimental Biology in Warsaw and is now head of basic research at the United Nations Educational, Scientific and Cultural Organization (UNESCO) in Paris, prices are

higher in Poland because fewer suppliers are actively competing to sell products there.

It is a badly kept secret that many scientists bring in supplies themselves from the West. "I know many good Polish scientists who always return from the United States with suitcases full of supplies and scientific devices," says Potempa. "I take things from there myself, every time I travel."

Researchers in eight eastern European countries can look forward to some price relief when they join the European Union (EU) on 1 May. After then, they can complain to the European Commission if they are restricted from shopping around other EU countries to get the best price.

For the rest of the world, there is little relief in sight. The situation continues to frustrate researchers. Stevens Rehen, a neuroscientist with positions at both the Scripps Research Institute at La Jolla, California, and the Federal University of Rio de Janeiro, for example, would like to set up a new lab in Brazil. But with pricing as it stands, he says, "it is almost impossible to establish research in Brazil". ■

Full results of the survey are at

♦ www.nature.com/nature/pricesurvey

Bioterror tester kits trouble federal agencies



Test cases: emergency workers say they need kits to make faster checks for dangerous pathogens.

Erika Check, Washington

The US Department of Homeland Security (DHS) is running a \$1.5-million programme to evaluate hand-held kits used by emergency workers to test for biological hazards in possible terrorism situations.

The programme is the latest of several major federal efforts to evaluate the kits over the past two years. The kits are used by first responders — workers who are first on the scene in any emergency — who need to determine quickly whether a suspicious substance contains a pathogen such as anthrax. The programme will set operating standards and test which devices meet them.

Such kits are controversial, and battles over their use illustrate the technical challenges and difficulties in coordinating federal agencies that the US government faces as it tightens homeland security.

The government first focused on the kits in late 2001, after a terrorist mailed anthrax

spores around the country. These high-profile attacks were followed by tens of thousands of 'powder calls' — most of which were false alarms. To distinguish fake calls from real ones quickly, emergency workers began using hand-held detectors. But federal officials were unsure about the accuracy of the devices, so the FBI and the Centers for Disease Control and Prevention in Atlanta, Georgia, both ran tests on them.

The results alarmed the US government's science advisers. On 19 July 2002, John Marburger, head of the White House Office of Science and Technology Policy (OSTP), announced in a memorandum that commercial hand-held detectors were plagued by technical problems and advised first responders and government officials not to use them. Marburger said the kits were prone to delivering false positive results, which could result in costly and frightening quarantines and city shut-downs.

This announcement angered some emergency workers, who say they know the limitations of the kits and combine them with other techniques to rule out false positives. They also say that the government's advice — that first responders rely only on laboratory tests — is not practical.

"The hand-helds can rule out a lot of things so, potentially, we do not have to quarantine people on a train for 72 hours while we wait for an answer," says John Eversole, chair of the International Association of Fire Chiefs' Committee on Hazardous Materials.

The manufacturers of the devices were also upset, because they say the first evaluation of hand-held kits took place behind closed doors, without their input.

Last year, the DHS and the OSTP began new evaluations. They set up a committee to oversee the tests under the auspices of the Association of Analytical Communities, based in Gaithersburg, Maryland. The committee comprises academics, government officials and industry representatives.

But controversy remains as manufacturers are worried that the tests will be biased by the OSTP's earlier conclusions and say the government is out of touch with everyday situations faced by first responders. "There is significant tension between the guy in the field and the bureaucracy," notes William Nelson, chief executive of Tetracore, a biotech company in Gaithersburg.

And scientists testing the devices remain concerned. "This product is too limited," says Vincent Vilker, head of the biotechnology division at the National Institute of Standards and Technology, which is one of the agencies evaluating the results of the tests.

But with first responders clamouring for help, DHS and OSTP officials say they must provide guidance soon — at least until better kits come along. The new evaluations are expected to be available later this year. ■

Rebelling scientists welcome left's landslide in France

Declan Butler, Paris

French researchers protesting against the government over deep cuts to science funding emerged strengthened after last weekend's regional elections, which saw the ruling conservatives wiped off the map by a left-wing alliance.

The Socialist party, with their Green and Communist allies, took 50% of the vote — a score not seen since François Mitterrand's presidential victory in 1981. The Union pour un Mouvement Populaire, the party created by President Jacques Chirac in 2002 to unite the right, took 37% leaving it with only

Alsace; previously, the party held 14 of France's 22 regions.

Mass demonstrations by researchers in the run up to the election contributed to the vote against Prime Minister Jean-Pierre Raffarin. The revolt hit a public nerve, symbolizing what was wrong with the administration.

"The unprecedented protests by researchers played a significant part in the results of the elections, something never seen in recent history," says Vincent Courtillot, a geophysicist at the Paris Geophysical Institute, on the Jussieu campus in Paris. Courtillot should know. He was principal

adviser to the minister for national education and research Claude Allègre who, following a smaller scientists' uprising in 2000, was simply sacked (see *Nature* 404, 421; 2000).

Courtillot is now optimistic about progress. Never before have researchers been so united on the shape of needed reforms, he points out (see *Nature* 428, 105; 2004). "A significant fraction of requests stand a better chance of being heard now," he says.

Chirac is expected to hold on to Raffarin, at least until June elections to the European Parliament, but to implement a thorough government reshuffle. ■



Defensive move: the US administration wants to enhance the capabilities of cruise missiles.

Australia considers revised scheme for young researchers

Carina Dennis, Sydney

Australia is seeking to revamp a training scheme for young scientists that has been irking the nation's universities.

The change is proposed in reviews of Australian research, released on 24 March, which also called for A\$500 million (US\$375 million) to be spent on boosting collaboration between institutions, and for a top-level advisory council to be set up to guide national science policy.

The three reviews looked at university research funding, research infrastructure and collaborations between universities and public research agencies. The government is expected to take them into account in a long-term budget plan to be released next month.

The recommendations include an overhaul of the nation's A\$540 million per year research training scheme (RTS), which funds postgraduate students. Universities have complained that the complex formula it uses to allocate funding penalizes universities whose students complete their courses early.

Last year, for example, the University of Melbourne began legal action against the federal government for revenue it claims it lost because of the scheme. "The university estimates it has lost more than A\$10 million over the past three years," says Kwong Lee Dow, its vice-chancellor.

Federal science minister Brendan Nelson said after the reviews were released that the government would be making some changes to the RTS.

The reviews, conducted by scientists and government officials, also called for A\$500 million to be spent over ten years to drive collaborations among universities, research agencies and industry.

"I am fairly optimistic that there will be funding for this, and I'm hopeful it will be on top of existing budgets," says Kurt Lambeck, a geophysicist at the Australian National University in Canberra.

The reviews further suggest creating a strategic research council to make decisions about national research policy. This is seen as a mechanism for getting the main research agencies to work more closely together. But some people worry that it could just create more bureaucracy.

"There is no real indication of what the structure or function of the council would be," says Snow Barlow, an environmental physiologist at the University of Melbourne and president of the Federation of Australian Scientific and Technological Societies.

Democrats slam Bush plan for fresh nuclear weapons

Geoff Brumfiel, Washington

The Bush administration is pushing ahead with plans to research, and perhaps build, new kinds of nuclear weapons, angering critics who say that this will encourage global proliferation.

The programme was outlined by energy secretary Spencer Abraham and Linton Brooks, head of the National Nuclear Security Administration, at congressional hearings over the past two weeks. It would see some \$500 million spent on advanced concepts for nuclear weapons over the next five years. This would pay for research and development of a nuclear-tipped cruise missile with advanced navigation capabilities and a weapon known as a bunker buster, which could strike targets buried deep underground.

Leading Democrats lined up to denounce the plan, saying that the energy department has other nuclear-weapons priorities that seem to be suffering at the expense of the new programme. "You're rushing ahead with new nuclear weapons including mini-nukes and nuclear bunker busters," said Senator Edward Kennedy (Democrat, Massachusetts). "But you're cutting funds for nuclear security."

Since the Bush administration unveiled its nuclear doctrine in 2001, officials have advocated the development of low-yield nuclear weapons that might be used against hardened or deeply buried targets without causing massive civilian casualties. They have also argued that working on such weapons will strengthen the skills at the nuclear-weapons labs — Lawrence Livermore in California, and Los Alamos and Sandia in New Mexico (see *Nature* 415, 945–946; 2002). But critics counter that the availability of the weapons will increase the possibility that they will be used, by the United States or by others.

The president's budget for 2005 proposes that funding for research on "advanced concepts" should increase by 50% to \$9 million. According to Abraham, much of that would go towards developing a cruise missile that could be retargeted after launch and would incorporate new safeguards against accidental detonation. Funding for research into a hardened warhead designed to destroy deeply buried targets would also increase from \$7 million to \$27 million.

But Democrats are angry at a five-year plan for these projects, contained in a five-year forecast of the budget, which suggests spending roughly \$450 million for the development of the bunker-busting weapon between 2006 and 2009.

Abraham told the hearings that, for the moment, the weapons labs are carrying out only paper studies and have no plans to build or test new weapons. "All we are offering Congress is a cost assessment of what the programmes might be," he said.

But some observers say that the five-year plan exposes the administration's real intentions. "Having budget numbers that extend through the development phase makes it harder for the administration to convince senators that this is just a research project," says Michael Levi, a physicist at the Brookings Institution, a think-tank in Washington DC. "This indicates to me that the administration clearly intends to move forward," adds Daryl Kimball, executive director of the Washington-based Arms Control Association, which advocates non-proliferation.

Last year, congressional opposition halved the administration's funding proposal for research into the bunker buster. This year, with an election at stake, it can expect an even tougher fight.

Founder bows out as genomic firm slashes workforce

Meredith Wadman, Washington

Human Genome Sciences (HGS), a biotechnology company that epitomized the promise of genomic medicine, is laying off 200 of its 1,000 employees and parting ways with its charismatic chief executive, William Haseltine.

Haseltine says that he was the one who initiated his retirement from the company, which is based in Rockville, Maryland. The former Harvard AIDS researcher, who founded the company in 1992, says his skills are not the commercial ones now needed to make drugs available for patients. HGS board members confirm that Haseltine was not ousted.

The company also announced on 25 March that it was jettisoning several drugs in early-stage development to focus on its five most promising candidates, which include LymphoStat-B for rheumatoid arthritis and systemic lupus erythematosus — now in phase II trials.

Some industry observers say the changes in the company, which follow a cut of 7% of employees last December, are normal for a biotech company as its drugs near commercialization. "It is just a natural part of the evolution of a company in drug development," says Robert Eaton, president of MdBio, a non-profit organization that supports biotechnology in Maryland.

Others disagree. "Things are getting tough at HGS and Haseltine's leaving the ship when it looks as though there are reefs in the distance," says Alan Walton at venture-capital firm Oxford Bioscience Partners in Westport, Connecticut, who helped launch HGS.

HGS has \$1.3 billion in cash, but still lacks a drug close to commercial launch. Its shares, of which Haseltine owns 6 million, are worth \$12, compared with \$232 at their peak in March 2000.

Haseltine's departure echoes that of his long-time rival Craig Venter who left Celera Genomics, also based in Rockville, in January 2002. Both men are aggressive scientific entrepreneurs who seduced Wall Street with high profit expectations that are so far unfulfilled. Their rivalry developed during a contentious partnership between HGS and The Institute for Genomic Research, a non-profit research organization founded by Venter, which ended in 1997.

The HGS board plans to hire a chief executive by the autumn and already has a list of candidates. Haseltine will retire once his successor is named. ■



Bottom of the heap: dung beetles and other lowly life-forms struggle to mount conservation agendas.

Biologists get ball rolling to aid neglected invertebrates

Helen Pearson, New York

Conservation biologists are drawing up a comprehensive plan for the preservation of invertebrates — populous but overlooked inhabitants of the animal world.

Ranging from beetles to jellyfish, invertebrates make up 90–95% of Earth's animal species and form food for much of the remainder. Although the exact numbers under threat are unknown, habitat destruction, pollution, invasive species and climate change are driving numerous species towards extinction.

Many researchers are aware of the plight of invertebrates, but such organisms have struggled to gain a foothold in conservation agendas. For example, only 189 species of invertebrate are listed as threatened or endangered under the US Endangered Species Act, less than 40% of all the animal species listed. And with very few additional species now being added to that list (see *Nature* 426, 592; 2003), the prospect of adding uncharismatic flies and worms seems to be remote.

Now 300 invertebrate researchers, policy-makers, conservation project managers, and representatives of non-governmental organizations have met to discuss the issues facing invertebrate conservation. The Expanding the Ark symposium, which took place on 25 and 26 March at the American Museum of Natural History in New York, was extended when some 25 participants decided to form a consortium that will act as an advocacy group for these neglected species.

Experts at the meeting were concerned that creepy-crawlies are not as winsome as tigers and bears — yet public concern can be critical for a species to earn legal protection.

Researchers are also troubled that the vast

majority of invertebrate species are unknown, unnamed or uncharacterized — and without this information it is hard to make a case for their protection. "We have not delivered the data to make invertebrate conservation possible," says Sacha Spector, who helped to organize the meeting and studies invertebrates at the American Museum of Natural History.

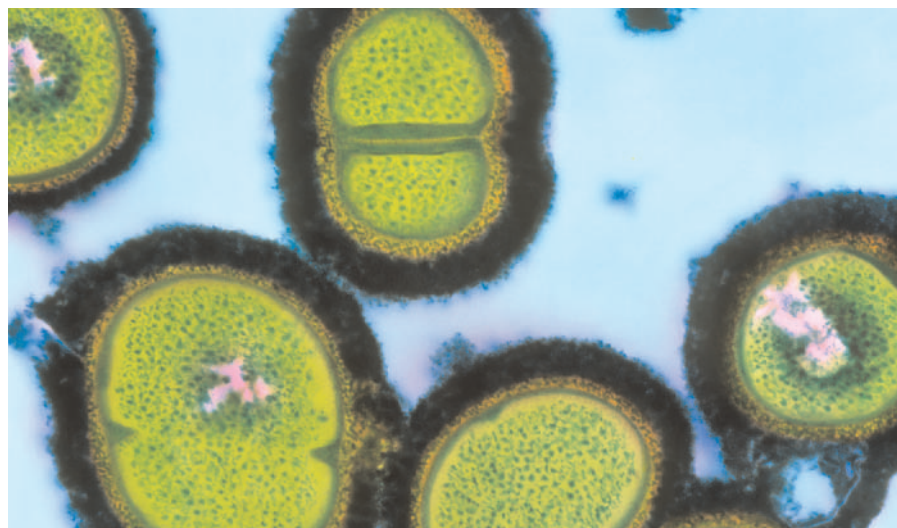
Spector says the consortium plans to produce two documents this year. The first will draw together published data on the decline in invertebrate species in order to advertise their plight. The second will draw up a set of recommendations for reversing the decline, which will be aimed at scientists, conservation groups, educators and policy-makers.

Details are still being discussed but the recommendations will include better ways to collect and access biodiversity data, such as collaboration with amateur conservationists. "We'll be very interested in that," says Dan Ashe, science adviser to the director of the US Fish and Wildlife Service, which enforces the Endangered Species Act.

The paper is also likely to advocate economic evaluations. According to a study presented at the meeting by entomologist John Losey of Cornell University in Ithaca, New York, insects alone provide services worth \$65 billion a year in the United States: decomposing dung, pollinating crops, and eating pests.

Whether or not the paper really makes an impact, delegates at the meeting say it will show that the invertebrate conservation community is getting its act together. "This is the first time that something like this has started to gel," says Scott Hoffman Black, executive director of the Xerces Society, a conservation group in Portland, Oregon. ■

Defence work sheds light on hospital bacteria



A light-sensitive method allows rapid detection of deadly methicillin-resistant *Staphylococcus aureus*.

Laura Nelson, London

Researchers have developed a test that will speed up the identification of antibiotic-resistant bacteria in hospitals, allowing quicker treatment for patients and curbing the spread of infections.

The test tracks down methicillin-resistant *Staphylococcus aureus* (MRSA) bacteria by tagging them with luminescent enzymes and measuring the light they give off. It uses technology first developed at the UK Ministry of Defence bioweapons laboratory at Porton Down in Wiltshire to monitor airborne bacteria for bioweapon detection.

"We have cut down the time needed for the diagnosis of MRSA," says Bill Mullen, chief executive of Acolyte Biomedica, a company based at Porton Down science park that has patented the test. Acolyte Biomedica has a licence agreement that gives it access to the defence lab's work.

Worldwide infections of MRSA have steadily escalated since it was first detected in the 1960s. In Britain, where the problem has caused widespread public alarm, MRSA costs the National Health Service £1 billion (US\$1.8 billion) each year, according to the National Audit Office, which audits the

accounts of government departments and agencies.

MRSA is usually carried asymptomatically on the skin, but causes infection when it enters wounds or affects weak immune systems. It can be fatal because its resistance to antibiotics makes it difficult to treat.

In the test, samples are taken from patients' nasal fluids and the swabs placed in liquid containing magnetic beads coated in antibodies that stick to the bacteria. These are then removed from the liquid using another magnet.

If MRSA cells are present in the sample, the beads will then be covered in them. To test this, the liquid coating the beads is placed in a dish with the antibiotic methicillin. This kills all other bacteria, leaving only the MRSA to divide and multiply.

After a few hours, the firefly enzyme luciferase, which reacts with adenosine triphosphate (ATP) in living cells to produce light, is added to the mix. The culture will glow if MRSA is present. Another compound, adenosine diphosphate, which amplifies the amount of ATP produced, is also added, increasing the test's sensitivity 100-fold. Other tests for MRSA take days, rather than hours.

Mullen is pleased the military work will help to address a public-health problem. "We are turning swords into ploughshares," he says. He adds that the test will be tried out in hospitals from May and will be available for general use by early next year. ■

Letters to the Editor deliver fitting April message

Alison Abbott, Munich

Long ago, the residents of the Polish village of Chelm reputedly decided that the Moon was more important than the Sun, because it provided light when it was really needed — at night.

This outlook apparently rules the roost at the Chelm Institute in Orange County, California. Each year, the *Journal of Biological Rhythms*, a circadian-biology journal, publishes unusual letters from the institute's researchers in its April issue.

Last year, for example, Nikolai Ivanovich Lobachevsky wrote in from the institute complaining of the "sappy sermons you call editorials"; and giving readers advice on how to exploit the peer-review system to see off their competitors. In a previous year, the Chelm institute's more idealistic M. Pupique argued for a lottery system for selecting papers for publication or making academic appointments, to overthrow what he termed the "hegemony of the academic ruling class".

But the journal's editor Martin Zatz, a circadian biologist at the National Institutes of Health in Bethesda, Maryland, has now come clean and revealed that he writes the letters himself. They are April Fool's Day spoofs, intended to stimulate thoughts on important issues in science that Zatz considers worthy of a wider airing.

"I want to make the journal more interesting," he explains, "and some of the issues that interest me relate more to the human-nature side of science — issues that are difficult to address in editorials."

The letters attributed to Lobachevsky and Pupique allow Zatz to address concerns about the imperfections of peer review and proposed alternatives to it, Zatz says.

In this month's spoof, Quincy Adams Wagstaff, director of the Chelm Institute Omega Multi-Center Institute for

Co-operative Research in the 21st Century, boasts of his alleged discovery of the "Mother of all Clock Genes". Wagstaff's self-aggrandizing text praises the teamwork that

led to the discovery — but it soon becomes apparent that the bombastic director has no intention of rewarding the rest of the team with either co-authorship or a share in a pending patent.

Some of the references in the spoof letters may pass many readers by. Not everyone, for example, will recall that Lobachevsky was a nineteenth-century Russian mathematician,

whose name was appropriated by Tom Lehrer in a 1960s song about plagiarism, or that Wagstaff was an anarchic college president in the Marx brothers' film *Horse Feathers*. Zatz is bemused by such shortcomings. "Few understand irony any more," he complains. ■



Retribution denied to creationist suing arXiv over religious bias

Washington A lawsuit that accused arXiv, a preprint server for physics and astronomy papers, of religious discrimination has been thrown out of court on a technicality.

Robert Gentry, a Seventh Day Adventist and former Oak Ridge National Laboratory scientist, filed a suit in 2002 against arXiv after administrators removed his ten papers detailing an alternative hypothesis to the Big Bang. Gentry, a creationist, claimed that the open-forum preprint server had discriminated against him on religious grounds. The website's curators said that Gentry's papers had been removed because he lacked proper academic credentials (see *Nature* **420**, 597; 2002).

The suit was dismissed from a Tennessee court on 23 March because Gentry failed to show that the server, or its operators — Cornell University in Ithaca, New York, the National Science Foundation and Los Alamos National Laboratory in New Mexico — had sufficient presence in the state to merit legal action.

Paul Ginsparg, a physicist at Cornell who founded arXiv in 1991, described the ordeal as “irksome” but says that the suit has led to

new policies at the website. Since January, anyone wishing to post to the website has had to win a referral from a current member. “We are trying to facilitate communication within professional communities,” Ginsparg says. “The endorsement system makes that process more transparent and maintainable.”

Mars Express confirms presence of methane

London The detection of tiny amounts of methane in the martian atmosphere has been confirmed by equipment on board the European Space Agency's orbiting Mars Express.

The gas was first detected in trace concentrations by Earth-based telescopes in Hawaii and Chile last year. Vittorio Formisano, principal investigator operating the craft's Planetary Fourier Spectrometer, told *Nature* this week that his team had confirmed the presence of the gas.

As methane has a residence time in the atmosphere of just a few hundred years before breaking down, it must be being constantly replenished. A geological source is possible, especially as Mars was still volcanically active just 100 million years ago. Another theory is that ancient but now extinct life could have left buried methane

hydrate deposits from which the gas is slowly leaching.

Unfamiliar diet blamed for elk deaths

San Diego The mystery surrounding the deaths of more than 300 Rocky Mountain elk in Wyoming has been attributed to overindulgence in a plant not found in their normal habitat.

Since early February, veterinarians and biologists have been investigating what caused apparently healthy elk to become immobilized, fall to the ground and die. Common ailments, including chronic wasting disease, were all ruled out.

In a study earlier this month, three test elk were fed a suspect lichen (*Xanthoparmelia chlorochroa*). “Two of



Lichen are suspected of causing elk deaths.

M. HAMBLIN/OSF

the elk are already down with identical symptoms," said state veterinarian Walter Cook last week.

Autopsies on the test elk show damage to major muscles that is similar to that seen in the elk found dead. Blood and tissue analysis are under way to identify the culprit toxin, which Cook believes may be usnic acid. Drought is thought to have pushed the elk north of their normal range and into areas where the lichen is abundant.

Interior department rejoins the web pending appeal

Washington The US Department of the Interior regained access to the Internet on 25 March after ten days out in the cold. The hiatus left the whole department, which manages conservation projects across the United States, unable to communicate effectively with researchers internally and externally.

Web access was suspended by a federal judge on 15 March following allegations that the department had failed to protect trust-fund accounts from hackers.

This is the third time since 2001 that security concerns have interrupted the department's Internet connection. It stopped scientists from submitting requests for proposals, reviewing government projects or looking at job postings.

NASA jet shatters speed record

San Diego An experimental aircraft last weekend flew at seven times the speed of sound, smashing the speed record for a jet plane. NASA's X-43A craft flew under its own power above the Pacific Ocean for ten seconds on 27 March, reaching speeds of more than two kilometres per second.

The X-43A is the culmination of NASA's Hyper-X programme to develop jet engines that can propel planes at 'hypersonic' speeds — faster than Mach 5. Engineers at the space agency hope that this may lead to a new generation of reusable space vehicles that rely on jet engines rather than rockets.



J. ROSS/NASA

Following an emergency request, Internet access was restored pending the department's appeal against the judge's ruling.

Website launched for gene-therapy trials

Washington US authorities have revamped the way that adverse events on human gene-therapy trials are reported. The National Institutes of Health (NIH) has previously been criticized for not ensuring that researchers report such events immediately (see *Nature* 403, 237; 2000).

Details of adverse events will now be submitted via a website, the Genetic Modification Clinical Research Information System. The NIH, which runs the site with the Food and Drug Administration, hopes this will speed up submissions. A request by the NIH in 2000 for information on adverse events in gene-therapy trials, after the death of a patient on a study by the University of Pennsylvania, revealed that only 6% of events were reported at the time they occurred.

The website also includes a database of US trials, and details of the therapy used.

► www.gemcris.od.nih.gov

In the know

In big groups and companies, it's hard to track who knows what. So how can scientists share information and prevent work being duplicated? Philip Ball investigates one solution: knowledge-management software.

It's a common frustration in scientific life: you have a seemingly simple query, but you aren't sure who can answer it. A basic experimental procedure, for example, might already have been mastered by someone else in the same building or company. Talking to that person could save you weeks of work — but how do you find them?

Researchers in small and specialized groups can simply ask around. But what about those in pharmaceutical companies that employ thousands of people around the world? Or scientists in multicentre collaborations? Asking individual colleagues would be far too time-consuming, and a single bulk e-mail would annoy many recipients.

Knowledge-management systems may be the answer. The software builds up a picture of who knows what in an organization, and uses the information to connect queries with answers. After a series of false starts, such systems have had some success in the pharmaceutical industry. And with electronic networking now embedded in scientific life, the infrastructure is there to implement knowledge management in new ways.

"Does the 25-person company need a knowledge-management system?" asks Eytan Adar, an information scientist at Hewlett-Packard Laboratories in Palo Alto, California. "Probably not. But would a 10,000-person pharmaceutical company, with researchers on both sides of the Atlantic, like to keep track of lab work to avoid duplicated experiments? There the answer is more clear."

Knowledge management is an organizational memory bank. Every problem that a staff member solves, from testing a candidate drug molecule to curing a glitch in computer



Someone, somewhere probably knows the solution to your problem — but how do you find them?

code, adds to this memory. And if this treasure trove is accessible to others, the problem need only be solved once.

The longer a journal editor is in the job, for example, the more he or she relies on experience when choosing referees. There will be times when an editor knows that two researchers have fallen out, and should not review each others' papers. But how do others find that information? And how might this knowledge be preserved when the editor leaves?

Electronic memories

Keeping a record of everything is impractical, and useful information would be swamped in a sea of trivia. Often it is a chance remark at the coffee machine that becomes a crucial tip-off — it is peer-to-peer communication, not managerial fiat, that ensures important knowledge is passed on.

The first attempts at creating an electronic organizational memory sought to store and mine electronic data. One of the most influential was Answer Garden¹, developed in the late 1990s by Mark Ackerman, a computer scientist at the University of Michigan in Ann Arbor. This expert-finding system uses a

database of employees' expertise, as volunteered by them, to find the right person to answer a question. Users are presented with a set of frequently asked questions (FAQs) that might avoid the need to consult another employee. If this fails to satisfy, they can be connected with an expert through e-mail, and the answer provided is added to the FAQ stockpile.

Other systems based on matching users with experts appeared around the same time, but promise gave way to disillusionment. Simon Masterton, a knowledge-management expert at the food and chemicals company Unilever, based in Port Sunlight in northern England, summarized the problem last year. "Very often there is a positive initial response to a new knowledge-management system," he wrote in a paper with colleague Stuart Watt, now at Robert Gordon University in Aberdeen, Scotland². "Many people will try it out a few times. Unfortunately, more often than not the use of the system will begin to tail off pretty rapidly."

That's just what Ackerman found when academic groups and computer companies trialled Answer Garden³. "Things look great in theory," says one user. "Then when you get

W. TAUFIC/CORBIS

to the specifics, it looks like more effort than it's worth. You really don't have the time."

There were other, more subtle, problems. Users found that useful knowledge could not always be decontextualized and stored. Some people were unwilling to share. And Answer Garden could not provide the incentives needed to compensate for the time and effort of sharing⁴.

Tools for knowledge management in science also flowered briefly around the time of the first business-software projects. In the early 1990s, Eric Mandel of the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, worked with Ackerman to develop a knowledge-management system for astronomers called ASSIST⁵. Astronomers often analyse many types of data with different, incompatible software packages. ASSIST embeds such software in a single electronic environment, containing advice and tutorials written and updated by users. Some users call it a living cookbook.

Astronomy cookbook

Within the first four months of its 1992 release, ASSIST was retrieved by more than 40 astronomical institutes, and was used on NASA projects such as the Compton Gamma Ray Observatory and the Advanced X-ray Astrophysics Facility (now called Chandra). Nevertheless, it did not catch on. Mandel says that it worked well, but lacked backing. "We really needed one large astronomical project to use ASSIST so that its truly innovative capabilities could be demonstrated in practice. But we never got one," he says. The NASA projects made limited use of ASSIST, and no one was willing to develop the tutorials needed to test the software, Mandel adds.

But scientific knowledge management is not dead. In fact, it is thriving, Ackerman says, but in a different guise. He points to arXiv, the physics preprint server hosted by Cornell University in Ithaca, New York, as a simple implementation of peer-to-peer sharing. "Scientists are not only lead users of knowledge-management techniques," says Ackerman, "they are inventing new techniques that are



Well connected: Mark Ackerman has devised programs to help researchers share their expertise.

then incorporated into business practices."

New knowledge-management software has also emerged in recent years. Employees of the pharmaceutical company Aventis are using one new system, called KnowledgeMail. In 2001, for example, a researcher at DG Thrombotic Diseases and Degenerative Joint Diseases, a research division of Aventis based in Frankfurt, Germany, needed to figure out how to culture and sort the macrophages — a type of white blood cell — he was working with. He suspected the information was out there, but knew it could take weeks of searching to track it down.

With about 5,000 research staff and 75,000 employees, there was a good chance that someone in Aventis had already solved these problems. KnowledgeMail quickly found two researchers in the company's US division at Bridgewater, Massachusetts, who could supply the information.

Mail mining

The software, produced by Tacit of Palo Alto, California, develops an expertise profile of users based on words and phrases extracted from their e-mails. Crudely put, if one researcher gets a lot of e-mails about macrophages, the system assumes that they are an expert on the topic and sends appropriate questions their way. This allows expert profiles to be generated with little effort by users. Tacit's software has also been used by drug company AstraZeneca and technology firm Lockheed Martin; and KnowledgeMail has now

evolved into a product called ActiveNet.

Over a three-month trial in 2001, Aventis estimates that KnowledgeMail saved the company 7.8 person months. More than 80% of the 435 trial users in Germany, France and the United States asked for the software to be retained. Aventis is now extending use of the system, and other pharmaceutical companies are exploring similar packages.

But systems such as KnowledgeMail are not problem-free. "The issue with creating detailed profiles is the need for a guarantee that user privacy would be maintained," says Adar. Allowing users access to their profiles seems one way of addressing such concerns.

Adar and his colleagues have designed Social Harvesting of Community Knowledge⁶, software that builds up expertise profiles using e-mail scans, information on web pages visited and documents viewed. Users can also list their interests and skills in a self-declared profile, and can edit their profile or prevent certain sources from contributing to it. This facility was important to trial users, even though it may compromise the need to maintain unbiased profiles — people's self-image is often far from realistic.

If systems such as Adar's can continue to iron out such problems, new knowledge-management software could overcome the barriers that defeated early packages. Within certain research environments, they have the potential to cut down on tedious literature searches. And if they can do so without invading users' privacy or placing too much of a burden on their time, knowledge-management software could join the suite of computer programs that we cannot imagine living without. ■

Philip Ball is a consultant editor for Nature.

1. Ackerman, M. S. *ACM Trans. Inf. Syst.* **16**, 203–224 (1998).
2. Masterton, S. & Watt, S. *Inf. Syst. Front.* **2**, 299–315 (2000).
3. Ackerman, M. S. *Inf. Technol. People* **9**, 10–24 (1996).
4. Ackerman, M. S. & Halverson, C. in *Social Capital and Information Technology* (eds Wulf, V. & Huysman, M.) (MIT Press, Cambridge MA, 2003).
5. Ackerman, M. S. & Mandel, E. *J. Organ. Comput. Electr. Comm.* **9**, 105–127 (1999).
6. Adar, E., Lukose, R., Sengupta, C., Tyler, J. & Good, N. *Inf. Syst. Front.* **5**, 15–28 (2003).

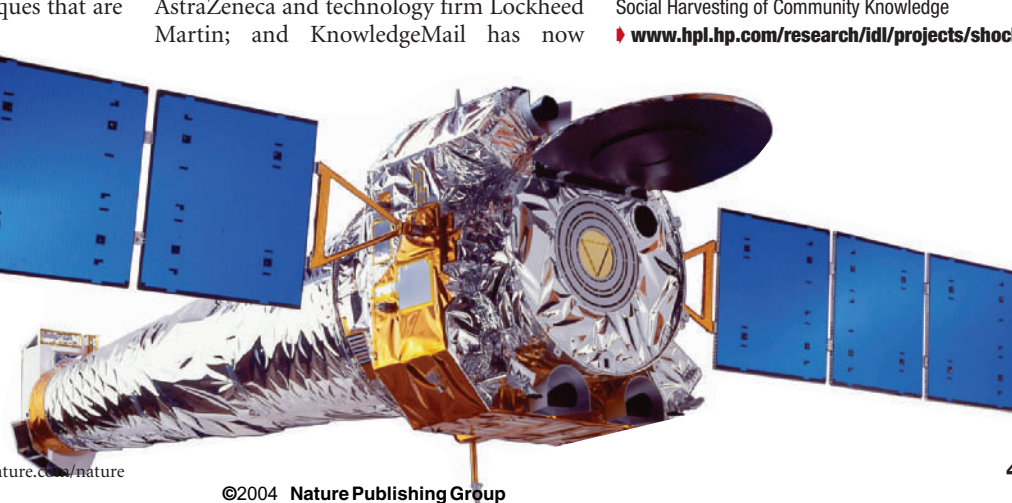
KnowledgeMail

► www.mindsharesolutions.co.uk/products

Social Harvesting of Community Knowledge

► www.hpl.hp.com/research/idl/projects/shock

Off line: the team behind the Chandra X-ray telescope lost its early enthusiasm for knowledge management.





The Renaissance rat

Thanks to the availability of its genome sequence, and the promise of new genetically engineered strains, the rat is restoring its reputation as researchers' favourite lab animal. Alison Abbott hails a remarkable rodent.

Rats tend to provoke strong reactions. Many people, repulsed by the rodents' tendency to feast on refuse and to spread disease, hate them. But physiologists, pharmacologists, toxicologists and pathologists are among the rat's biggest fans. For them, the rat is the most important experimental animal — and one for which they can't hide their affection.

"I love them," admits John Fozard, a pharmacologist with the drug company Novartis in Basel, Switzerland. "They're wonderful little beasts: easy to house, nourish and handle, and scientifically very reliable." Today's lab rats have a rich scientific pedigree (see 'Roll of honour', opposite), and their value is now being enhanced by the availability of the rat genome sequence, described in this issue of *Nature*¹.

Not before time, say the rat's scientific cheerleaders, who have met regularly since the 1970s under the unofficial banner of the 'Rat Pack'. For in recent years, superior genetic techniques used on mice — and information they have generated — have seen the rat's smaller cousin rise to prominence.

The rat's contributions go to the heart of biomedical research — quite literally, as it is

the most significant model for human cardiovascular disease. It's a similar story for diabetes and arthritis, and many behavioural disorders. When it comes to major health issues, rats are similar enough to humans to provide useful experimental data. And pharmaceutical companies use the animals for a large proportion of their mandatory toxicity testing.

Researchers particularly appreciate the rat's relatively generous proportions, which make it easy to carry out detailed physiological measurements. Yet once the human genome project was nearing its end, it was the mouse that became the next mammal in line to have its genome sequenced².

It's a knockout

That decision was influenced by the views of geneticists, who have always favoured the faster-breeding mouse. But even those researchers who complained that little was known about mouse physiology conceded that mice had gained a formidable technical advantage in the form of gene 'knockout' technology. Pioneered in the late 1980s, this allowed strains to be produced that lacked a single, specific gene.

Since then, the technique has been

refined to allow genes to be 'knocked in' as well as out. Today, geneticists can even make 'conditional' mouse mutants, in which a gene can be switched on and off at will in different tissues. Largely as a result of these technical advances, mice became more attractive both for genetic studies and as experimental models of human disease. During the 1990s, the ratio of scientific publications based on the rat or on the mouse decreased from 2:1 to 3:2 (ref. 3).

Until very recently, it seemed that the rat would not allow itself to be genetically manipulated in the same precise way. Knock-out mice are made by performing targeted genetic modifications on mouse embryonic stem cells, which are then injected into mouse embryos. By crossing the resulting animals over several generations, you can derive a line of mice bearing the desired modification in all of their tissues.

For reasons that aren't completely understood, it has proved impossible to isolate stem cells from rat embryos. Over the past couple of years, scientists working with sheep and pigs managed to sidestep this problem in their mammals by using cloning^{4,5}. Instead of modifying embryonic stem cells, the researchers manipulated a

gene in the nucleus of adult cells, and then fused these cells with egg cells that had been stripped of their own chromosomes. The resulting clones bore the genetic modification made to the original cells.

But cloning proved exceptionally difficult in rats. At the first hint of physical disturbance, rat egg cells spontaneously activate, and then cannot be made to implant in the uterus and develop to term. This was overcome when a team led by Jean-Paul Renard of INRA, the French agricultural research agency, in Jouy-en-Josas, worked out how to delay egg activation using a chemical that inhibits an enzyme involved in the process. His team reported the successful production of rat clones last autumn⁶ — and genOway of Lyon, the company that owns the technology, says that it will offer licences to academics wanting to make knockout or knock-in rats.

Under development

Others researchers, including Philip Iannaccone at Northwestern University's Feinberg School of Medicine in Chicago, Illinois, have recently succeeded in controlling the activation of rat eggs using different approaches, which they hope will also lead to successful cloning. Iannaccone wants to work with genetically manipulated rats to investigate a gene called *Gli1*, which is involved in a key developmental signalling pathway, and has also been associated with cancer in humans. Mice have not been so helpful to him, because knocking out the mouse version of the gene doesn't cause any noticeable biological changes.

It will take some time for the Rat Pack to catch up with their mouse colleagues, but the advent of genetically engineered cloned rats should help to level the playing field. And just as the mouse genome provided a host of new gene targets to knock in and out, so too will its rat counterpart. "Scientists are going to be able to choose the animals they want to work on based on biological rather than technical considerations," predicts Howard Jacob, a rat geneticist at the Medical College of Wisconsin in Milwaukee. Being able to compare both rat and mouse genomes with the corresponding human sequence will also provide a sound basis to relate the results of experiments conducted in rats and mice to human disease. "You can do the physiology in the rat and the genetics in the mouse," says Jacob.

Until now, much of the research done on rats has used outbred strains such as the Sprague-Dawley or the Wistar, which are relatively diverse genetically. But even before the advent of cloning, physiologists had access to some rats with defined genetic characteristics. Over the years, breeders have generated more than 300 inbred strains — all members of each are genetically identical. The Brown Norway, the strain whose genome has now been sequenced, has been the most popular. Other inbred strains model human disease.

Roll of honour

1621: Sexual confusion

Theophilus Müller and Johannes Faber of Italy's Accademia dei Lincei in Rome perform the first recorded rat dissection (right). Their pregnant wild specimen appears to have a uterus, penis and testes, and they describe it as a hermaphrodite. The penis was actually a clitoris, and the testes were vaginal glands.

Early nineteenth century: Into the lab

The brown rat becomes the first animal species to be domesticated for scientific use. Early physiological studies concentrate on the effects of food and oxygen deprivation.

1877: Well bred

Twenty-three years before the monk Gregor Mendel's work on inheritance comes to prominence, researchers carefully interbreed rats with different coat colours and note their results.

1900: Model behaviour

Rat intelligence and learning is tested in the 'Small maze', a labyrinth based on the maze at the former royal palace of Hampton Court in England. Mazes (below) and other behavioural tests are refined over the next century to study learning, memory and the brain.

1906: Setting the standard

Researchers at the Wistar Institute of Anatomy and Biology in Philadelphia interbreed albino rats to produce the first Wistar rats. More than half of all laboratory rats used today are thought to be direct descendants of these rodents.

1920: The big C

Rats fed on tapeworm eggs develop liver tumours within eight months, becoming the first rat model of cancer. Subsequently, researchers develop inbred strains that are prone to particular cancers. Fischer 344, for instance, spontaneously develops leukaemia and prostate cancer.



1963: Under pressure

Researchers inbreed Wistar rats and produce the Spontaneously Hypertensive rat. These animals are still used to test and develop drugs to treat hypertension, and to help find genes that regulate blood pressure.

1971: Short of breath

Rats are given small amounts of a protein such as egg albumin to stimulate their immune systems. When the treatment is repeated, their lungs become inflamed, mimicking human asthma.

1980: In the swim

The Morris water maze is developed to test rat learning and memory — the animals must learn to find a platform in a murky pool of water. Easy to use, the test is widely adopted to

study topics including stroke and neurodegenerative disease.

1980s onwards: Spare-part surgery

Rats have donated and received transplants of skin, heart, bone marrow, liver, small bowel, pancreas and brain tissue, allowing researchers to identify key molecules involved in rejection, and to combat this process.

2002: Remote control

Need a replacement for your robot? The radio-controlled rat could be the answer. A computer sends electrical signals to electrodes embedded in its brain, guiding the animal to turn left or right. Another signal stimulates the brain's reward centre — providing an incentive to follow these instructions.

Helen R. Pitcher

Rats of particular strains may develop metabolic diseases such as high blood pressure and diabetes — or even the rodent equivalent of depression.

And although knockout and knock-in rats are only now about to become available, it has for some time been possible to genetically engineer rats by the cruder technique of injecting a single gene into one of the unfused nuclei of a newly fertilized egg. The first transgenic rat was developed using this method in 1990 by Detlev Ganten at the Max Delbrück Center for Molecular Medicine in Berlin⁷. This rat contains an additional gene for the enzyme rennin, which triggers a sequence of events leading to rampant high blood pressure.

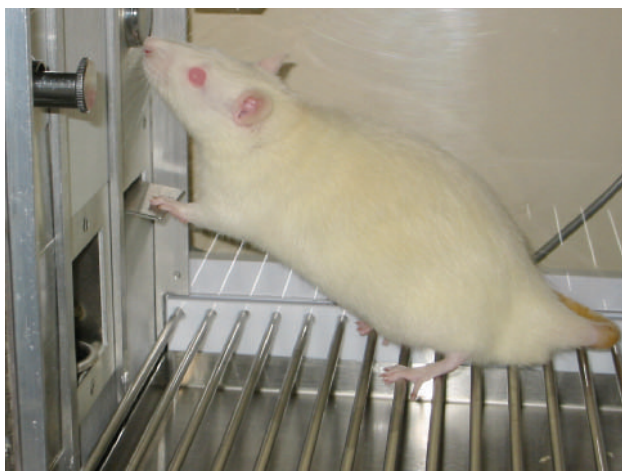
Bred for success

By laborious breeding programmes, scientists have also developed inbred strains that differ by a single chromosome, or just a small region of one chromosome. Called consomic and congenic strains, respectively, these rats can help to locate areas of the genome housing genes that increase the risk of succumbing to a complex disease such as diabetes. Jacob, for example, is leading a programme called PhysGen, funded by the US National Institutes of Health (NIH), which is developing consomic strains systematically and documenting their detailed physiology and pathology.

Jacob is also treating rat sperm with chemical mutagens in an attempt to generate rats with mutations in genes that change their susceptibility to disease. Other researchers have already used this method to generate a couple of rat strains in which genes that suppress breast cancer are inactivated⁸. The genomics company Ingenium Pharmaceuticals in Martinsried, Germany, began a similar project a year ago. It is making profitable use of the rat genome sequence, which has been available online for 18 months. "It was a big step forward that the rat genome, so well annotated, was made public," says Reinhard Sedlmeier, Ingenium's director of genomics.

In addition to supporting the development of new rat strains, the NIH is funding efforts to get them into the hands of researchers across the United States. In 2001, the agency began providing \$1.5 million per year for the Rat Resource and Research Center at the University of Missouri in Columbia. The centre currently maintains and distributes about 30 different strains — a figure that is expected to grow rapidly. "The number of useful rat strains is exploding," says John Critser, the centre's principal investigator.

That is great news for physiologists who have flirted with the mouse in the past few years but who ultimately were left dissatis-



Model students: rats are ideal for behavioural studies because they learn tricks, such as pressing a lever to get alcohol, more readily than mice.

fied. Many had miniaturized equipment for measuring heart rate and blood pressure, or for imaging of blood circulation, but were disappointed with the results. "Size matters when it comes to physiological measurements," says Tim Aitman, who heads the Physiological Genomics and Medicine group at the UK Medical Research Council's Clinical Sciences Centre in London. Apart from the difficulty of working with miniaturized equipment, says Aitman, analysing the smaller quantities of blood that can be drawn from mice gives bigger margins of error.

Meeting of minds

Michael Spedding, head of research at the drug company Servier in Neuilly-sur-Seine, France, points to the superiority of rats for modelling human psychiatric stress. In stressed rats, the areas of the brain that change size are the same as those thought to be affected by stress in people. Mouse brains are just too small for researchers to measure such changes accurately.

What's more, even the most skilled experimentalists cannot get around the fact that, physiologically, rats are more similar to us than mice. For example, the mouse heart flutters away at about 600 beats per minute, whereas the rat heart beats at less than two-thirds that rate, closer to the average human resting rate of 70 beats per minute.

Behavioural scientists also prefer the

larger and less manic rat. Mice are difficult to work with because they scurry around so much; they are also slow and inflexible learners. A classic experimental set-up for researchers interested in addiction, for instance, involves teaching rats to press a lever to get a reward such as heroin. "Rats are smart; they know what they are doing," says Rainer Spanagel, who studies addiction at the Central Institute of Mental Health in Mannheim, Germany. "Mice just don't have a plan." They rush about pressing the lever randomly, he complains. "You need twice as many mice as rats to get statistically sound data from an experiment." And neurosurgery, which can be used to investigate the

brain structures involved in behaviour, is much more demanding in mice.

As anyone who has kept rodents as pets will tell you, rats are also less aggressive than mice — which is a boon for the technicians who have to handle the animals. This is due to the species' social structure in the wild: mice live in groups where one highly aggressive alpha male monopolizes the females. Rats, by comparison, conform to the 1960s ideal of blissed-out free love, with widespread promiscuity and low levels of aggression.

The rat's natural advantages as an experimental animal, combined with a wealth of new genetic information, should lead to a surge of interest among biomedical researchers. A measure of this was seen at the biennial Cold Spring Harbor Laboratory Meeting on Rat Genomics and Models in New York state last December, which attracted 25% more abstracts and a third more attendees than the previous meeting. GenOway, meanwhile, is now trying to put together a Rat Consortium to bring together animal breeders, academics and experts from the pharmaceutical industry to discuss their priorities in the generation of genetically modified rat models. The Rat Pack, it seems, is coming of age. ■

Alison Abbott is Nature's senior European correspondent.

1. Rat Genome Sequencing Project Consortium *Nature* **428**, 493–521 (2004).
2. Mouse Genome Sequencing Consortium *Nature* **420**, 520–562 (2002).
3. Hedrich, H. J. in *Handbook of Experimental Animals: The Laboratory Rat* (ed. Krinke, G. J.) 3–16 (Academic Press, New York, 2000).
4. McCreath, K. J. et al. *Nature* **405**, 1066–1069 (2000).
5. Phelps, C. J. et al. *Science* **299**, 411–414 (2003).
6. Zhou, Q. et al. *Science* **302**, 1179 (2003).
7. Mullins, J. J., Peters, J. & Ganten, D. *Nature* **344**, 541–544 (1990).
8. Zan, Y. et al. *Nature Biotechnol.* **21**, 645–651 (2003).

Rat genome

♦ www.hgsc.bcm.tmc.edu/projects/rat

PhysGen

♦ pga.mcw.edu

Rat Resource and Research Center

♦ www.nrrrc.missouri.edu



Ralph, the first cloned rat, was born last year.

Ancient frog could spearhead conservation efforts

Newly discovered, this unique creature is already under threat from dam projects.

Sir—The discovery of the purple or pignose frog *Nasikabatrachus sahyadrensis* was recently reported by two groups (see *Nature* **425**, 711–714; 2003, and *Current Science* **86**, 211–216; 2004). The sole member of the family Nasikabatrachidae, this creature evolved some 50 million years earlier than any other known frog in India.

Remarkably, the frog was well known to local people, suggesting that science does not pay enough attention to local knowledge.

This 'new' discovery was the culmination of many field studies on herpetofauna that gained momentum during the 1990s in the

Western Ghats. In a News and Views article, "The coelacanth of frogs" (*Nature* **425**, 669–670; 2003), Blair Hedges decries the policy carried out by some governments of restricting foreigners' access to biodiversity hot spots, and considers this an impediment to the discovery of rare and endangered species.

However, both of the groups who arrived at these results relied minimally on Western expertise and infrastructure. There is still a need for dedicated molecular-genetics facilities for ecology and evolutionary biology in developing countries. Building local capacity and

infrastructure may be the best way to accelerate the discovery of new species in remote areas, and this is where responsible scientists must focus their efforts.

The purple frog also has important conservation implications. Inundation of large chunks of valley forests in the Western Ghats by dam projects could spell disaster to this ancient amphibian and many other endemic taxa. *Nasikabatrachus* can be used as a flagship to promote the conservation of important habitats in the region.

Ramesh K. Aggarwal

Centre For Cellular And Molecular Biology,
Uppal Road, Tarnaka, Hyderabad 500 007, India

Diversity project takes time but reaps rewards

Sir—Your News Feature on the HapMap project (*Nature* **425**, 758–759; 2003) characterizes the Human Genome Diversity Project (HGDP) as an "ill-fated exercise" that "never got off the ground". This is not accurate. The birth of the HGDP required time and perseverance because of misunderstandings during its development, which led to a 1997 review by the US National Research Council. This reviewing committee was persuaded of the scientific merit of human genome diversity projects, while also recognizing the ethical and legal challenges implicit in studies of human variation.

Their faith in the HGDP has been rewarded. Since April 2002 a collection of more than 1,000 DNA samples from 51 populations representing most of the world's genome variation has been available to non-profit research laboratories through a collaboration between the HGDP and the Fondation Jean Dausset-CEPH in Paris (see www.cephb.fr/HGDP-CEPH-Panel).

All samples used for this resource were collected with proper informed consent; the privacy of the persons volunteering these samples remains protected. Since 1997, the HGDP has developed and promoted the best legal practices available for studying indigenous populations.

Within months of its public release, the HGDP-CEPH resource became the basis of the most comprehensive description of the genetic structure of human populations ever undertaken (N. A. Rosenberg *et al.* *Science* **298**, 2381–2385; 2002). This publication, which was not focused on any clinical outcome, was selected as the best biomedical publication of 2002 by *The Lancet*.

An enlarged HGDP collection would enhance research on the history and structure of human populations from all parts of the world, allowing researchers to characterize the global variation in any genetic sequence of interest. This work is useful to both evolutionary and medical studies. Identifying the genetic bases of complex diseases (the focus of the HapMap project) and of variation in response to environmental exposures and to drugs requires an understanding of genomic variation among unaffected individuals (controls) of similar ancestry to affected individuals.

Defining appropriate control populations for modern genomic analysis is still in its infancy. It can be helped by continuing political and financial support for HGDP resources.

Luigi Luca Cavalli-Sforza

Department of Genetics,
Stanford University School of Medicine,
Stanford, California 94305, USA

Small change

Sir—Changes to the way in which the Medical Research Council (MRC) distributes funds in the United Kingdom are indeed welcome, as is any increase in the medical research budget (*Nature* **427**, 665; 2004). However, to put the latter point into perspective, the entire MRC budget is not much more than the increase of US\$729 million proposed by President George Bush for the US National Institutes of Health budget in fiscal year 2005. On this point, the best one can say is that a very small pie is now merely small.

Stephen E. Moss

Division of Cell Biology, Institute of Ophthalmology,
University College London, Bath Street,
London EC1V 9EL, UK

Irrelevance, buzzwords, coffee: it's the virtual PI

Sir—Several graduate students would like to suggest a modest proposal to the authors of the recent article "Functional genomic hypothesis generation and experimentation by a robot scientist" (*Nature* **427**, 247–252; 2004).

We've been doing a little research of our own and have generated an exciting hypothesis: we can build a Virtual Principal Investigator.

The software platform will consist of a tiny amount of background knowledge about a completely different system from the one being actively investigated, a buzzword-invention engine, coffee-consumption apparatus, a grant-writing node and a cat-o'-nine-tails that will tie everything together.

For much less money than an actual investigator, we created a machine that resides in an office, drinking coffee at a steady rate. At bi-monthly intervals, the Virtual Principal Investigator spits out grants that are derivatives of highly successful grants from the past. Between these time points, it will play an endless loop of inspirational messages and test the effectiveness of the whip as a method of controlling subservient organisms.

Initial tests of our Virtual Principal Investigator have revealed no differences from our existing model of an investigative scientist, but we have yet to perform an experiment that was not naive or random. We look forward to discussing our results with you.

**Several non-robot scientists
(names supplied)**

Department of Biological Chemistry and Molecular
Pharmacology, Harvard Medical School,
240 Longwood Avenue, Boston,
Massachusetts 02115, USA

Power to the people

An analysis of the problems caused by an increasing energy demand.

Energy at the Crossroads: Global Perspectives and Uncertainties

by Vaclav Smil

MIT Press: 2003. 448 pp. \$34.95, £22.95

Michael Grubb

Energy and the environment are huge and complex topics, and this is Vaclav Smil's synthesis of a lifetime's study in these areas. It emerges as the most sober, thorough and thoughtful integrated text on energy available, and it embodies core facts and some fundamental truths that any analyst of energy issues should ponder.

It is rarely good practice to review a book chapter by chapter, but each of the six in Smil's opus merits specific mention. "Long-term trends and achievements" provides an overview of the extraordinary development of the world's energy system and the innovations that revolutionized the twentieth century. Smil doesn't shy away from recognizing the blessings of modern energy consumption, and the way these drove an explosive growth in demand. After some mind-numbing statistics on the expansion of world energy during the twentieth century, he refers to the drudgery of former village life and notes that "it is not easy to convey the liberating power of electrification ... no other kind of energy affords such instant and effortless access." Electrification also revolutionized manufacturing even more than steam engines did. Did you know that electric motors consume two-thirds of all electricity in the United States?

In "Energy linkages", Smil explores how energy statistics relate to the world, with close attention to energy demand. Energy intensities — the ratio of countries' energy use to their gross domestic product — "confirm and reinforce some widely held snapshot notions about the economic fortunes of nations". Yet he also warns against over-reliance on aggregated numbers, and notes huge measurement problems (energy intensities for China differ by a factor of six according to whether currency exchange uses conventional rates or those corrected by purchasing-power parity — "both are clearly wrong").

Smil goes on to delineate six key variables that influence national energy intensity. He lays out the volatile history of energy prices and the distinction between market and real costs of energy. He also charts the relationship between energy and quality of life: on all the key indicators, he concludes, quality of life is highly correlated with energy consumption during basic economic development, but is almost completely uncorrelated once countries are industrialized.



Generation ex? Solar Two, the world's largest solar tower, stopped producing electricity in 1999.

His third chapter, "Against forecasting", is by far the most comprehensive and damning demolition of energy forecasting I have seen. It is sobering to be reminded that only thirty years ago there were serious predictions that nuclear power would account for more than 90% of new construction by the start of this century, with equally far-fetched, unfulfilled and extraordinarily expensive programmes for coal, including the ill-fated US Synthetic Fuels Corporation.

The "eternally receding goal" of fusion power makes me similarly uneasy about the reviving fad for governments to reach for technology fixes to ease resurgent energy concerns. The human, institutional and policy element is vital too in energy demand — illustrated perhaps most starkly by the fact that, during the last six years of Mao's rule, Chinese energy intensity rose by 34%, then fell under Deng Xiaoping's liberation by a staggering 60% from 1980 to 1995, wrong-footing the analysts.

Buried within the chapter is another indication of why Smil deserves to be taken so seriously. A chart of the range of energy forecasts made in 1983 for global demand in the year 2000 shows that almost all the big institutional forecasts were far too high, that 'soft energy' optimist Amory Lovins was far too low, and that Smil was accurate to within 3%. But he modestly passes this off as chance and goes on to demolish both the history of energy forecasting and the pillars that support it. The two big persistent errors, he says, have been the "spell cast by the mood of the moment ... and infatuation with novelties and simple, magical solutions through tech-

nical fixes". Smil concludes that what we need in the face of such intrinsic uncertainties are more, serious "normative scenarios, outlining what should happen rather than what is likely to happen" — futures to guide us, not forecasts to fool us.

The chapter on "Fossil fuel futures" should also be compulsory reading for the confused. In particular, I have never seen a fuller or a fairer account of the perpetual debate about oil and gas resource depletion. The pessimistic school is pitched against the professional optimists, who can reasonably claim to have history on their side in predicting that we will continue to find more oil. My own view is that the pessimists may be closer to the truth on conventional oil than they are on gas, but on fossil fuels overall the debate misses the important point. There is plenty of energy out there, both fossil and non-fossil, but getting it will require huge new investment — most of which is still targeted towards getting at carbon more cheaply, rather than developing alternatives.

In addressing non-fossil energies, Smil's experience and ecological professionalism come to the fore. Renewable energy offers credible alternatives but this is not an easy solution. Smil is particularly sceptical about harnessing even more of the world's net primary product for biomass energy, and uses basic biochemistry to explain that ethanol production based on yeast-mediated fermentation of starchy crops can use only "a small fraction of the phytomass present in relatively expensive feedstocks". This highlights how the book's depth is both its weakness and strength: the heavy technical content makes

for hard reading but, when translated, it explains the fundamental reason why the ethanol subsidies for which the rich world's farmers lobby will do nothing for energy supply — unlike some of the other biomass energy routes.

Wind is perhaps the most interesting renewable: "generating 20% of US electricity demand would require less than 1% of land in the contiguous US, of which less than 5% would actually be taken up by the turbines themselves." The constraints will be those of systems and time, not natural resources. This is particularly true of the hydrogen economy, where Smil offers well-judged cautions and questions about current enthusiasms.

In his concluding chapter on possible futures, Smil emphasizes the complexities and timescales of change, which leads him to be generally conservative in the short term but with a broad long-term vision of what might be possible. This is true for both supply and demand: despite the book's emphasis on energy efficiency, Smil dismisses the fantasies of tenfold improvements being accomplished within a single generation.

Smil has the wisdom to frame the really big issues. He rejects the wishful thinking that the problems of energy are not real and the false hopes that they will go away or solve themselves. Truth matters, and is found neither in the specific scare stories or techno-optimism of many environmentally oriented writers, nor in the growth-blinded pronouncements of economic modellers and forecasters.

Yet Smil derives two big questions from the fundamentals of each of these world concerns: "If our actions were guided by the two greatest concerns a sapient terrestrial civilisation can have — for the integrity of the biosphere and for the dignity of human life — then it would be inescapable to ask the two most fascinating questions in energy studies: what is the maximum global TPES [primary energy supply] compatible with the perpetuation of vital biospheric services, and what is the minimum per capita energy use needed for a decent quality of life?" Smil says these questions are rarely asked, not only because they are extraordinarily difficult to answer, but because they demand clear moral commitment. Charting the energy horizons of environment and inequality, his conclusion boils down to Ghandi's famous dictum: the world has enough for everyone's need, but not enough for everyone's greed.

The policy implications outlined by Smil are not simple. How to promote efficiency and innovation without increasing consumption? How to reduce energy use among those who already overconsume fuel and boost it among those who don't? "When facing so many uncertainties, we should pursue any effective means that bring us closer to those goals," with "no categorical exclusion of certain ingredients (such as nuclear or big dams), no inflexible insistence on what is best."

This book should change the map by which we navigate the new energy century. In its intellectual content it is a great book, standing head and shoulders above most integrated writings on energy and environment. The complex, dense prose and tendency for rigorously balanced views may inhibit readers, but the book's underlying themes could and should prove the book's legacy. For example, Smil discusses the essential unpredictability and openness of energy futures, and provides a comprehensive critique of the notion that incessantly rising energy demand in industrialized countries is associated with increasing welfare — though doing anything about it will not be easy. The book has gone straight to the top of my reading recommendations for my students. And for the serious technical analyst of energy systems, there is nothing better. ■

Michael Grubb is professor of Climate Change and Energy Policy, Imperial College London, Exhibition Road, London SW7 2BP, UK.

Autistic genius?

Autism and Creativity: Is There a Link between Autism in Men and Exceptional Ability?

by Michael Fitzgerald
Brunner-Routledge: 2003. 304 pp.
£29.99, \$47.95

Autism: Mind and Brain

edited by Uta Frith and Elisabeth Hill
Oxford University Press: 2003. 298 pp.
£70, \$110 (hbk); £29.95, \$49.50 (pbk)

Allan Snyder

We tend to see only the whole — it takes insights from abnormal minds to appreciate how the brain assembles the parts. So could studies of autism reveal insights into creativity? It seems unlikely. The classical portrait of autism includes low intelligence, significant learning disabilities, memory by rote, literalness and a rigid insistence on sameness. Even autistic savants, known for their extraordinary mental feats, are not creative. Rather, they adopt a form of mimicry, probably due to privileged access to non-conscious processes. Beate Hermelin, an expert on autism, says that no savant will discover a new mathematical theorem, initiate a novel stylistic movement, or render a revealing interpretation of a Beethoven piano sonata.

But as these two superb new books demonstrate, our view of autism is radically transforming. It now embraces the classical picture of severe mental impairment at one end of the spectrum and possibly Nobel-prizewinning creative genius at the other. Both extremes have in common certain core autistic features, such as preoccupation with detail, obsessional interests and difficulties in understanding another person's perspective.



Did autism help shape the thinking of Indian mathematician Srinivasa Ramanujan?

Michael Fitzgerald, author of *Autism and Creativity*, says that some aspects of high-functioning autism and Asperger's syndrome enhance creativity. Because these developmental disorders are mainly genetic in origin and largely affect men, he believes that creativity in a broader sense is predominantly the result of genetic rather than environmental factors: "The view that geniuses began their lives made from the same material as the rest of us ... is false." He then engages in retrospective diagnosis to support such claims, declaring that several individuals "with creativity of genius proportions" fit the high end of the autistic spectrum. These include Isaac Newton, the philosopher Ludwig Wittgenstein, mathematician Srinivasa Ramanujan, Lewis Carroll, the poet W. B. Yeats, and politicians Keith Joseph and Eamon de Valera. Apparently Hitler too had autistic traits.

Fitzgerald's thesis is not new. Hans Asperger spoke of "autistic intelligence" as being intelligence of "true creativity", adding "it seems that for success in science or art a dash of autism is essential." Oliver Sacks suggested that Wittgenstein had autistic traits. So too did Einstein, van Gogh and possibly Bill Gates, according to Temple Grandin, who is herself autistic. Asperger even noted that the autistic mind is an extreme variant of male intelligence. Despite these earlier revelations, Fitzgerald's tantalizing book is a must read, as are Simon Baron-Cohen's brilliant contributions to this area, such as *The Essential Difference* (Perseus, 2003).

The fact that genius can fall within the autistic spectrum challenges our deepest notions of creativity. Are there two different routes to creativity: normal and autistic? The normal mind is good at recognizing the gist of something but poor at recalling details. This, I believe, is because the brain forms concepts or mental models that encapsulate the familiar. Concepts impart automatic judgments and confer intuition, but hide details from conscious awareness. As a result, we see the whole but not the parts. In contrast, the

Science in culture

Medical models

The German Hygiene Museum in Dresden is reopening its doors.

Alison Abbott

Last year an undetonated Second World War bomb was discovered near the German Hygiene Museum in Dresden. It was a reminder, if any were needed, of the dramatic history of the museum, which this week reopens its valuable permanent collection, 14 years after German reunification.

Most of the bombs did their job in 1945, however — the museum was 80% destroyed. Gone, too, was the Nazi ideology that had perverted its original, progressive, aim: enlightened health education.

By the start of the twentieth century, the infectious basis of common diseases was understood and, in the absence of antibiotics, the emphasis was on avoiding infection. Interest was high — some 30 nations took part in the First International Hygiene Exhibition held in Dresden in 1911.

The exhibition, which attracted more than five million visitors, was so successful that a permanent German Hygiene Museum was founded in Dresden the following year, moving into its permanent home — an imposing building in the Bauhaus tradition — in 1930.

Workshops sprang up to manufacture anatomical models and teaching aids of ever greater ingenuity and no small artistic value. These included moulages, or wax models of healthy or diseased body parts, transparent human organs produced by a new method developed by the Leipzig doctor Werner Spalteholz, and didactic wall charts. Most famously, Franz Tschakert prepared plastic-coated anatomical statues known as *Gläserne Menschen* (transparent people), complete with real or artificial skeleton, wax organs and thousands of metres of painted wire representing blood vessels and nerves.

The first Transparent Man was unveiled at the opening of the museum's new building in 1930. He stood on a plinth in an apse-like recess, a setting that no doubt added to the spiritual atmosphere he exuded with his exultantly outstretched arms.



It's clear to see: the Transparent Woman reveals her internal organs at the Hygiene Museum.

But three years later the atmosphere changed. Extending the general enthusiasm for hygiene to racial hygiene and eugenics, Nazi Germany subverted the museum for propaganda. In 1933 the Nazis also sacked the museum's Jewish employees and destroyed the frescoes in its restaurant painted by expressionist Otto Dix, as they considered him a degenerate artist.

The museum somehow maintained its professional reputation to the extent that its rebuilding became a priority for the Russian occupying army shortly after the bombing, and was completed within a year. The museum became a flagship of the new Communist state of East Germany, where it continued its official mandate to educate the public

about health issues. It spoke to the wider world — albeit with forked tongue — of East Germany's strengths in technology and education. Despite shortages of materials, the manufacture, sale and export of its anatomical models grew. It produced the first black anatomical models for export to Africa, and developed special wax for moulages exported to hot climates. It sold more than a hundred transparent men and women, extending its repertoire to transparent cows and horses.

But reunification put an end to the museum's official purpose — West Germany already had a government office for health education.

The permanent exhibition is part of the museum's development of a fresh identity. The first of the four new exhibition rooms is dedicated to the Transparent Man, which remains a metaphor for enlightenment and transparency about the past, as well as allowing us to see tissues and organs. Instrumentation on show includes a first-generation X-ray machine. And there is a range of historical anatomical models, including the first Transparent Woman and a more recent model whose organs light up when buttons are pressed. The other rooms feature life and death, eating and drinking, and sexuality. Three additional rooms, dedicated to movement and the cardiovascular system, neuroscience, and skin, will open next year.

The museum's curators have eschewed recent trends towards amusement-park styles of presentation. This is a place to learn, and the museum's important collection is central. Unobtrusive computer terminals provide further information, which can be e-mailed to visitors' home computers. The displays' strong, simple and light design — the ultimate in German chic — creates an effect that is modern and learned, calm and uncluttered, despite the wealth of objects. Its elegance is a tribute to its early, avant-garde history.

Alison Abbott is Nature's senior European correspondent.

autistic mind is literal and sees more of the parts than the whole. An impairment in the process of concept formation denies intuition, but allows access to details that are normally non-conscious. Consequently, the autistic mind must build logically from the parts to what is intuitive to a normal mind.

The autistic mind seems to be suited to working algorithmically within a closed system of specified rules. In contrast, the normal mind can make unexpected connections between seemingly disparate systems — often by breaking the rules of each when taken in isolation, but not as an ensemble. In other words, a normal mind invents entire new systems rather than finding novelty within a previously prescribed space. Is it possible to have the best of both worlds? Could certain psychopathologies inadver-

tently plunge someone into a temporary state of autism, allowing them to see the parts normally denied to conscious awareness?

To gain deeper perspective on such issues requires information from diverse research. The second book under review, edited by Uta Frith and Elisabeth Hill, gives a valuable update in 13 insightful chapters, written by authorities on the subject. I especially enjoyed the editors' two chapters and those by the groups of Baron-Cohen, Happé and Schultz. Although tilted to the specialist, this excellent book portrays a panoramic view of autism. It is loaded with all kinds of goodies: autism is no longer a rare disease; it can be associated with congenital blindness; people with autism have difficulty recognizing faces; they show a strong desire to systematize; and on average their brains are larger

and heavier than normal brains from around the age of 2–4 years (but probably not as adults). Movement disturbances may play a role in autism — a reduction in facial expressions may reflect problems with the underlying social brain network.

We are told that there is still no unifying theory of autism. But I suggest that a failure in the process of concept formation and its associated top-down inhibition of the parts that make up the whole may offer a mechanism that could unite the current descriptive theories. Concepts order the world internally. Without them, order must be imposed externally, hence the setting up of rigid routines. ■

Allan Snyder is director of the Centre for the Mind, a joint venture of the Australian National University, Canberra, ACT 0200, and the University of Sydney, Sydney, New South Wales 2006, Australia.

Turning down, but not off

Neuroprotection requires a paradigm shift in drug development.

Stuart A. Lipton

The industrialized world faces an unusual predicament. The demographic shift towards an ageing population, coupled with the prevalence of Alzheimer's disease and vascular dementia (both of which cause cognitive decline), has led economists to estimate that by 2050, the entire economy of the industrialized world could be consumed by the costs of caring for the sick and elderly. What can be done to abate these dreaded neurodegenerative diseases and their socioeconomic effects?

Until recently, only symptomatic approaches were available — for example, cholinergic drugs for Alzheimer's disease, which mildly enhance memory by increasing levels of the neurotransmitter acetylcholine. But the concept of 'brain protection' raises hope that neurons can be preserved from the ravages of neurodegenerative insults.

In most neurodegenerative diseases, the brain is attacked and nerve cells are killed by a variety of overactive signalling pathways. These pathways are triggered by conditions such as oxidative or nitrosative stress, accumulation of aberrant proteins, and excessive activity in the brain of the neurotransmitter glutamate (excitotoxicity). Excitotoxic damage, a common final pathway contributing to most or all neurodegenerative disorders, is largely caused by overstimulation of NMDA-type glutamate receptors. This causes excessive influx of Ca^{2+} through the receptors' associated ion channels, resulting in detrimental enzymatic reactions and generation of toxic oxygen and nitrogen free radicals.

These neurodegenerative processes could be potential drug targets for neuroprotection. For example, injurious agents or instigators of degeneration could be counteracted, as in the case of excessive amounts of glutamate, which excite neurons to death, or aberrant proteins (such as mutant parkin protein in juvenile Parkinson's disease) that clog the brain's detoxification and degradation systems. Other examples include the depression of 'death' signalling pathways by inhibiting enzymes such as caspases; restoration of pathways that degrade aberrant proteins through the proteosomal and lysosomal systems; or enhancement of survival pathways, with either anti-cell-death proteins or neurotrophic factors that prolong neuronal life. Cell-based approaches could also be advanced with attempts to ameliorate inflammatory cell activities, and stem-cell replacement therapy to provide new brain cells.

But there is a paradox. Many of these potential targets have normal functions,



which only become deleterious to the nervous system when in excess or under maladaptive conditions. For example, physiological NMDA-type glutamate receptor activity is essential for normal neuronal development, communication between neurons and memory formation. Neuroprotective agents that work by high-affinity binding to these receptors block all activity, and these drugs produce unacceptable side effects including hallucinations, drowsiness and coma. Similarly, microglia and astrocytes nurture and protect the neurons they surround, and only release toxic substances when they are inappropriately perturbed. We cannot simply shut off these cell processes and molecules without compromising normal or adaptive functions.

What is needed is a way to target neuroprotective drugs and cell-based therapies at abnormal functions in an appropriate spatial and temporal pattern for the disease, while sparing normal, physiological activity. But drug discovery in the big pharmaceutical companies involves high-affinity screening of target molecules — hence these drugs work 'too well' and 'all the time'. If we use an analogy of the target molecule as a television set, these 'competitive' drugs battle one-on-one with the agonist at the on/off switch (the agonist binding site) and, when successful, will simply turn the set off. A neuroprotective agent that works by high-affinity binding to the receptor of a neurotransmitter will block all activity — normal and abnormal. In addition, such drugs will outcompete lower (physiological) levels of neurotransmitter more effectively than higher (pathological) levels, meaning that normal areas of the brain will be shut off

even before pathological areas are effectively protected. Thus, such drugs manifest unacceptable side effects by blocking normal physiological activity in all parts of the brain, even those not affected by the disease process.

What if the brain could be protected using drugs that do not bind very well under physiological conditions, but that under pathological conditions become selective for the target? Using the television analogy, such a drug would be equivalent to the volume control. For example, excitotoxicity could be prevented by turning down the excessive 'volume' of Ca^{2+} influx through the receptor's channel towards normal, thus avoiding the formation of free radicals. But if the drug binds with high affinity in the channel, it will accumulate there, and once again block normal function — turning the volume all the way down is as bad as turning off the on/off switch.

On the other hand, a clinically tolerated drug would block only excessive activity while relatively sparing normal function, simply by adjusting the volume towards normal levels. Such drugs are termed uncompetitive inhibitors — they work better when increasing levels of agonist are present, hence blocking excessive (pathological) receptor activity while sparing lower (physiological) activity. A key to this mechanism is not only drug selectivity for the target despite low affinity but, most importantly, the ability to come off the target relatively quickly, preventing accumulation and blocking of subsequent normal function. One such drug is memantine, which my colleagues and I discovered can preferentially block NMDA-glutamate receptor-associated ion channels when they are excessively open. Memantine has recently been approved for treatment of Alzheimer's disease in Europe and the United States.

By virtue of their relatively gentle binding, drugs of this type work best under pathological conditions, while exerting minimal effects on normal brain activity. This simple concept could be extended to other neuroprotective targets — even to other pharmaceuticals — and in my opinion will be significant in future drug design. ■

Stuart A. Lipton is in the Center for Neuroscience and Aging, The Burnham Institute, La Jolla, California 92037, USA.

FURTHER READING

- Chen, H.-S. V. *et al.* *J. Neurosci.* **12**, 4427–4436 (1992).
- Lipton, S. A. & Chen, H.-S. V. *Cell Death Diff.* **11**, 18–20 (2004).
- Lipton, S. A. & Rosenberg, P. A. *N. Engl. J. Med.* **330**, 613–622 (1994).
- Reisberg, B. *et al.* *N. Engl. J. Med.* **348**, 1333–1341 (2003).

Three's company

Kerstin Lindblad-Toh

Publication of the rat genome sequence will not only advance physiological studies in this paragon of laboratory animals, but also greatly enhance the power of comparative research into mammalian genomes.

On page 493 of this issue¹ there appears a description of the third mammalian genome to be sequenced and analysed — that of the Brown Norway rat. The excitement generated by this research has a number of roots, not least the fact that the rat is an important model for learning about human physiology and disease; knowing its genome sequence will greatly enhance our ability to associate genes and mutations with traits and diseases. Comparative mammalian genomics will also benefit, because having the rat, mouse² and human^{3,4} genome sequences will allow us to explore which characteristics are specific to rodents and which are shared by all mammals.

Rats (*Rattus norvegicus*; Fig. 1) were established as a model for learning about human physiology and disease in the early 1800s. In the 1900s, they ceded some of their popularity to mice, which are smaller, quicker to breed and easier to manipulate genetically. Still, the rat never fell out of favour in areas where its larger body size and physiological similarity to humans are important, including pharmacological studies — testing the effects and toxicity of drugs. (The history and importance of rats as lab animals is discussed in the News Feature on page 464 of this issue⁵.)

As with mice, there are several hundred inbred or partially inbred rat strains, which differ in many physiological measures, such as blood pressure or body weight, or in their propensity to diseases, such as diabetes mellitus or arthritis. Studies of these different strains have enabled researchers to associate many such traits with specific large chromosomal regions in the rat genome. Many of the characteristics are quantitative traits — they show a graded spectrum of variation from individual to individual — and have complex inheritance patterns, being controlled by more than one gene; they are detected in the genome as quantitative trait loci (see the rat genome database⁶). But the causative genes at these loci are, in many cases, unknown. Today's publication by the Rat Genome Sequencing Project Consortium of the complete sequence of the Brown Norway rat¹ will accelerate the pace of discovery of genes, alleles (gene variants) and mutations associated with these physiological properties and disorders. In fact, the identification of several disease-associated genes —



Figure 1 Model animal — the lab rat.

including ones involved in a cerebellar defect⁷ — has already benefited.

To take full advantage of the wealth of physiological variation among rat strains, researchers also need a map of the genetic variation. The new sequence will serve as the basis for this map, on which researchers will mark the sequence positions that vary between the Brown Norway rat and other strains. As the consortium discusses¹, a pilot project to identify variations in the coding portions of genes (single nucleotide polymorphisms) is ongoing. A continued effort would generate maps that show which portions of the genome are highly similar and which are quite different between strains — similar to the 'haplotype maps' that are under way for humans⁸ and mice⁹.

Beyond its value in unlocking the treasures of rat biology, analysis of the rat genome sequence will also help biologists to identify and understand human genes and gene regulation, and to study mammalian genome evolution. On the evolution front, the consortium's comparison¹ of the rat and human genomes reveals a strikingly large amount of genome shuffling. The previous comparison of the mouse and human genomes² noted that most genes have the same immediate neighbours in both species, although blocks of genes have been shuffled considerably. Including the rat genome confirms the suspicion that most of the 300 flips

and swaps between rodents and humans happened in the rodent lineage after it split from the common ancestor with humans, and that 50 of these flips are also different between rats and mice.

Similarly, the rate of change of individual bases seems to be much higher in the rodent than the human lineage, judging by how much variation is seen within sequence repeats that are assumed to have been present in the common ancestor. Together, these findings hint that the rodent genomes are more dynamic than the human genome, evolving roughly three times as fast. This raises the question of whether the mutation rate is different in rodents — or whether it is simply that smaller animals have a shorter generation time and hence go through more mutations in eggs and sperm (mutations that are therefore passed on to the next generation) in the same time frame¹⁰.

To translate findings from a model organism to human medicine, it is necessary to correlate genes and potential disease-associated mutations across species. So the broad similarity in the number, order and sequence of genes between rodents and humans is reassuring, as is the fact that as many as 90% of rat genes have matches in both humans and mice¹. This number is higher than the 80% reported when comparing mice and humans². The difference is probably due to a better cataloguing of

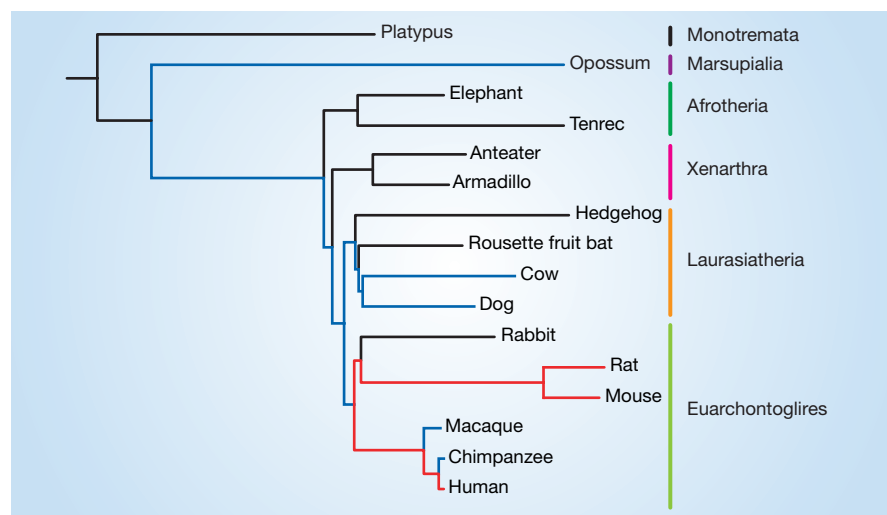


Figure 2 Mammalian evolution and genome sequencing. This evolutionary tree shows the position of mammals whose genomes have been sequenced and analysed (red) or are being sequenced (blue). Humans, mice and rats come from the same major clade (Euarchontoglires), dogs and cows from a further clade (Laurasiatheria). But two major clades of placental mammals remain unsampled (Xenarthra and Afrotheria). Note also the fast evolutionary rate (represented by a longer branch length) of rats and mice compared with humans.

genes and pseudogenes (copies of genes that are no longer functional) in the rat genome — possibly because of comparison across the three genomes.

The 10% of genes that are not present as one-to-one pairs between species belong to gene families that have expanded in number in one of the species. Many of these gene families are involved in olfaction, immunology or reproduction, and can easily be associated with biological features that make each species unique; for instance, humans rely less than rodents on their sense of smell, and so have fewer olfactory-receptor genes. Of particular interest for pharmacological studies are the detoxifying P450 genes, of which mice and rats have more than humans. So, it may be more difficult than thought to use the toxicity of drugs in rats as a guide to their toxicity in humans — because rats may be better at removing toxins from their system. Still, having the genome sequence makes it possible to assess whether the P450 genes acting on a specific drug are present in both rats and humans. Moreover, this gene number is more similar between rats and humans than between mice and humans, which could help explain rats' superiority over mice in pharmacology.

In the initial comparison between the mouse and human genomes², the high degree of sequence conservation found hinted that some 5% of each genome contains functional elements — genes, regulatory elements and so on. Including the rat genome has significantly increased the resolution of these analyses¹. The precision of gene prediction (annotation) will continue to increase with each vertebrate genome that is sequenced, and with the availability of other resources, such as the Mammalian Gene Collection¹¹,

that catalogue physical evidence of gene transcription. This is encouraging, because we are far from knowing every gene in the human genome, and even further from knowing how each gene is regulated. Using comparative sequence analysis, researchers can draw on the experiments performed by evolution, where functional elements are

conserved and neutral sequences — those sequences that have no particular function — change over time.

To accomplish this in the next few years, a working group¹² has been charged with developing a plan to select the appropriate species and sequencing strategy by which to identify all the functional elements in the human genome. Several other mammalian genomes (those of the chimpanzee, macaque, dog, cow and opossum; Fig. 2) are currently being sequenced, and when completed will further enhance the three mammalian genomes that we already have. The high quality of the rat genome sequence¹, and the speed with which it was produced, leads us to expect great things in the near future.

Kerstin Lindblad-Toh is at the Broad Institute, 320 Charles Street, Cambridge, Massachusetts 02141, USA.

e-mail: kersli@broad.mit.edu

1. Rat Genome Sequencing Project Consortium *Nature* **428**, 493–521 (2004).
2. Mouse Genome Sequencing Consortium *Nature* **420**, 520–562 (2002).
3. International Human Genome Sequencing Consortium *Nature* **409**, 860–921 (2001).
4. Venter, J. C. et al. *Science* **291**, 1304–1351 (2001).
5. Abbott, A. *Nature* **428**, 464–466 (2004).
6. <http://rgd.mcw.edu>
7. Kuramoto, T., Kuwamura, M. & Serikawa, T. *Brain Res. Mol. Brain Res.* **122**, 103–108 (2004).
8. The International HapMap Consortium *Nature* **426**, 789–796 (2003).
9. Wade, C. M. et al. *Nature* **420**, 574–578 (2002).
10. Weinreich, D. M. *J. Mol. Evol.* **52**, 40–50 (2001).
11. <http://mgc.nci.nih.gov>
12. <http://www.genome.gov/1000218>

Earth science

Inside history in depth

David Stevenson

The history of how Earth's interior evolved, and how it accounts for many aspects of our planet's behaviour, remains largely unwritten. Taking water into account could well help to explain a great deal more.

The basic divisions of Earth's internal structure (crust, mantle and core) have been known for a long time. But the evolutionary path that gave us this structure, and that provides the dynamics of plate tectonics, volcanism and magnetic-field generation, remains poorly understood. Why do we have plate tectonics? What is the nature and extent of melting deep within Earth? How does the core manage to keep generating such a richly complex magnetic field?

These were among the questions debated on the first day of a workshop* organized under the auspices of the CSEDI (Cooperative Study of the Earth's Deep Interior), a National Science Foundation programme that is now eight years old. From the talks and ensuing discussions among deep-Earth devotees, it is evident that we need a better

knowledge of the processes that govern deep-Earth history, and the material parameters that control those processes, before any kind of 'standard model' can be constructed — much though such a model may be desired by some participants.

A simple view of Earth's evolution invokes the first law of thermodynamics, estimates of radiogenic heat production — that produced by radioactivity — and some straightforward scaling arguments that emerge from our theoretical and experimental understanding of thermal convection. In this view, Earth started hot and cooled through geological time at a rate that closely mimics the decreasing rate at which the radiogenic heat production declines. This decreasing rate arises because the radiogenic heat sources have a variety of half-lives, and the more long-lived sources become increasingly important with time. Earth is

*CSEDI Science Plan Workshop. La Jolla, California, USA, 22–23 February 2004.

now presumably shifting to the predominance of ^{232}Th , the longest-lived of the sources. The other parameter of great importance in simple models is the temperature dependence of mantle viscosity, because this dictates the vigour of mantle convection and plate tectonics as the heat flow declines. In this approach, Earth's core has a heat flow that is strongly coupled to mantle evolution through the mantle's ability to cool and accept heat from the core.

Models of this kind are easy to construct and boringly monotonic. Furthermore, they cannot explain the widely accepted factor-of-two ratio for current Earth heat output to current radiogenic heat production. Our planet was more eventful than these simple models allow. Whereas Earth scientists have no desire to repeal the first law of thermodynamics, they are willing to challenge almost everything else. Recently, major disagreements have emerged in attempts to understand the energy budget of Earth's core, and there are still many uncertainties over how to incorporate the effects of plates, water, melting and layering into our picture of mantle circulation.

In the current debate about Earth's core, there is a contrast between low- and high-dissipation pictures of the dynamo responsible for the planet's magnetic field (P. Olson, Johns Hopkins Univ.). These respectively make different assumptions about the geometry of the magnetic fields and associated electrical currents in the core. There is much publication activity on this topic^{1–4}, but no consensus, because we don't know the full complexity of the field and currents involved. A new study⁵ invokes experimental dynamos as well as theoretical ideas, and suggests that dynamos may be less dissipative than some suppose, thereby making a dynamo easier to sustain. Uncertainties about the electrical and thermal conductivity of the core material, and its phase diagram, are also large enough to have a major effect on estimates of dissipation and convective vigour.

For decades, it has been popular to invoke Earth's solid inner core as one of the main contributors to the energy budget available to the dynamo. Latent heat and gravitational energy and buoyancy are released as the inner core grows and excludes some of the light-element components of the outer core (sulphur, oxygen and silicon, for example) from its crystalline structure. Standard evolutionary models have difficulty explaining how the inner core has existed for more than the past billion years or so, yet Earth's magnetic field has existed throughout most of geological time. There is no direct evidence on the age of the inner core, and the dynamo may operate without an inner core. Still, it would be surprising if it were a recent feature of Earth's structure. This is one of several reasons why some scientists wonder whether there is an additional energy source in the

core. A possible candidate is radioactive decay of ^{40}K . Although the amount of core radiogenic heating required may be modest compared with Earth's present total energy budget, that amount assumes increasing importance as one goes back in time because of ^{40}K 's relatively short half-life of around one billion years. It is not known whether even this modest amount of potassium in the core is assured by elemental partitioning between silicates and iron alloy at the time of core formation. Uranium is also a candidate, and additional non-radiogenic (gravitational) sources cannot be excluded.

Mantle convection driven by thermal buoyancy remains the agreed framework in which to understand plate tectonics and loss of Earth's heat. It is an unfortunate feature of simple models of convection that they can mimic many of the characteristics of plate tectonics, but cannot explain some essential features of plates. The danger of these simple pictures is that they may not provide an adequate predictive framework for how plate tectonics evolves through geological time. Some models^{6,7} suggest possible solutions, but the lack of agreement between these various approaches means that we are not close to a final resolution. Water may play a major role through its ability to lower the viscosity of the mantle and possibly the strength of plates. State-of-the-art approaches to elucidating mantle convection include ambitious attempts to describe mantle mixing and accommodate the geochemical constraints (L. Kellogg, Univ. California, Davis).

Partial melting of shallow mantle is the cause of oceanic crust formation, the relatively well-understood and dominant form of volcanism on Earth. But water and carbon dioxide have a major yet poorly understood effect on melting in the mantle (M. Hirschmann, Univ. Minnesota). We do not know the total amount of water in the mantle (although it is at least comparable to that in Earth's oceans), or the rate at which this reservoir is tapped and replenished. Nonetheless, the emergence of provocative ideas⁸ on the topic illustrate a growing willingness to tackle the central questions of Earth's interior ocean of water.

Remarkably, these seemingly disparate topics have a common thread. It seems likely that we will not understand the origin of Earth's magnetic field until we know how the mantle controls heat flow in the core. But we cannot understand the mantle side until we have a better understanding of plate tectonics. This may in turn depend on understanding Earth's water cycle. Could it be that magnetism, like life, depends on water? ■

David Stevenson is in the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA.
e-mail: djs@gps.caltech.edu

1. Buffett, B. A. *Geophys. Res. Lett.* **29**, 1566 (2002).
2. Gubbins, D. *et al.* *Geophys. J. Inter.* **155**, 609–622 (2003).
3. Labrosse, S. *Phys. Earth Planet. Inter.* **140**, 127–143 (2003).
4. Nimmo, F. *et al.* *Geophys. J. Inter.* **156**, 363–376 (2004).
5. Christensen, U. & Tilgner, A. *Nature* (in the press).
6. Korenaga, J. *Geophys. Res. Lett.* **30**, 1437 (2003).
7. Solomatov, V. S. *J. Geophys. Res. Solid Earth* **109**, B01412 (2004).
8. Bercovici, D. & Karato, S. *Nature* **425**, 39–44 (2003).

Entomology

Butterflies at that awkward age

Dick Vane-Wright

Beautifully preserved specimens of butterflies from the Caribbean, caught maybe in the act of egg-laying some 20 million years ago, provide welcome grist to the mill of debate about butterfly history.

Writing in *Proceedings of the Royal Society*, Hall and colleagues¹ describe a new fossil butterfly from examples preserved in amber from Dominica, on the island of Hispaniola. They have called the butterfly *Voltinia dramba*: about 20 million years old, it is the first adult butterfly to be formally described from Dominican or any other type of amber. Its discovery raises key issues about Caribbean biogeography, behavioural evolution (or lack of it), and the origin of butterflies.

The fossil is from the family Riodinidae (the metalmarks), and it belongs to a genus of nine living species now found only on the Latin American mainland (Fig. 1, overleaf). Over 1,000 species of riodinids are known from Central and South America, and in

some lowlands they form 20% of the local butterfly fauna. Yet of 300 butterfly species found in the West Indies today, only one is a riodinid, and that species is quite unrelated to the new fossil. So the first issue is how *V. dramba* reached Hispaniola.

With five known fossil examples, beautifully preserved, Hall *et al.*¹ are able to produce good evidence that one of the living species — *V. danforthi* from Mexico — is the closest relative of *V. dramba*. Given that the fossil specimens are 15–25 million years old, Hall *et al.* suggest that the best explanation for the butterfly's presence on Hispaniola is breakup of the 'proto-Greater-Antillean arc', which is postulated to have separated the Caribbean islands from Central America². They date this split, and that between *V. dramba* and

V. danforthi, at 40–50 million years ago. The implied survival time for the latter approaches an order of magnitude greater than the life-span³ of the average species, and Hall *et al.* suggest that *V. danforthi* qualifies as a 'living fossil'. If this is correct, an implication is that the basic ecology of *Voltinia* has not changed over this huge time span.

The ecological context here is that the caterpillars of some butterfly species feed on a wide variety of flowering plants. Others are very restricted, as for example in the close relationship between white butterflies and cabbages. This suggests that groups can become locked together by accidents of history, their only option thereafter being to co-evolve⁴. Alternatively, caterpillars and their butterflies can be fickle, chopping and changing from host to host as opportunity arises or necessity dictates. In evolutionary biology we must be alert to mere story-telling, selecting suitable facts to support whatever view of events we favour. But in this case, the authors' argument that the ecology of *Voltinia* has remained essentially the same for as long as 50 million years seems quite convincing.

Voltinia belongs to a small group of metalmarks in which the caterpillars have the unusual habit of feeding on bromeliads and orchids growing on tree trunks⁵. Both host-plant groups appear to have existed for at least 60 million years (R. Bateman, personal communication). Dominican amber is fossilized resin from an extinct leguminous tree, on which suitable host plants could have grown. Thus it does not seem surprising that a *Voltinia* would be the first adult butterfly found in amber. Moreover, all five known specimens of *V. damba* are females. Hall *et al.*¹ make the obvious suggestion that they

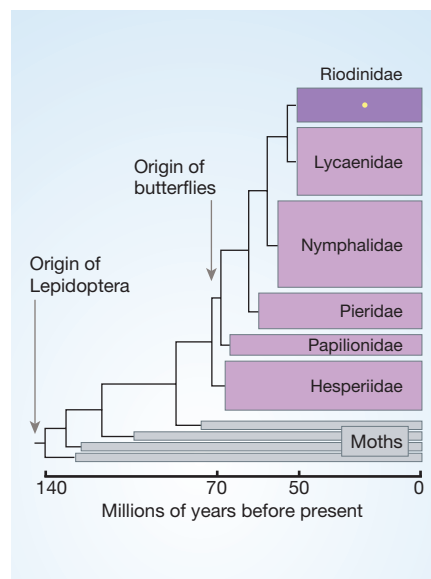


Figure 2 Relationships among the Lepidoptera — the butterflies (about 20,000 living species) and moths (ten times that number). Common names for the members of each butterfly family are metalmarks (Riodinidae); blues and coppers (Lycaenidae); admirals, heliconians, milkweeds and browns (Nymphalidae); whites and sulphurs (Pieridae); swallowtails (Papilionidae); skippers (Hesperidae). There are numerous separate lineages of moths. The *Voltinia damba* fossils described by Hall *et al.*¹ are about 20 million years old (yellow dot). The oldest known butterfly fossil, thought to be a skipper, is 52 million years old; and the oldest known fossil of a lepidopteran associated with a flowering plant shows the track of a caterpillar that fed inside a leaf 97 million years ago. (Data from refs 1, 8, 9, 12, 13.)

became trapped in resin while laying their eggs. But given a 1:1 sex ratio, the odds against getting an all-female sample of five would seem to be only 32:1 against. As this is a single, unrepeatable sample, is this just story-telling?

Male butterflies are usually encountered more often than females, even though captive breeding⁶ almost invariably reveals a true sex ratio of 1:1. For the group of five genera to which *Voltinia* belongs⁵, the apparent sex ratio of these butterflies in The Natural History Museum collection is 1.65:1 — a seemingly small bias in favour of males. Using this figure to derive an encounter frequency of 0.375 for females, instead of 0.5, gives a probability of over 100:1 against getting a sample of five female *Voltinia* at random. Although this calculation is open to objections, it nonetheless suggests that Hall *et al.* are right

to seek a nonrandom explanation — and a remarkable 40–50-million-year behavioural stasis in egg-laying behaviour and host-plant choice seems the obvious conclusion.

Turning to butterfly origins, there are good reasons to believe that the major diversification of the Lepidoptera, the butterflies and moths, is linked to the evolution of flowering plants, currently dated at a minimum age of about 140 million years. Many modern Lepidoptera families were established by the early Tertiary⁷, 60–70 million years ago, and butterflies branch off high in the 'crown' of the lepidopteran evolutionary tree⁸ (Fig. 2). This suggests a possible date of about 70 million years ago for the origin of butterflies.

A persistent belief of some lepidopterists is that many higher groups of butterflies evolved in the great southern continent of Gondwana⁹. This landmass started to fragment during the middle-late Jurassic, some 175–160 million years ago, with final break-up being complete by the early Tertiary. If butterflies were much older than this, the Gondwana hypothesis would be plausible; if much younger, then it would be implausible. An age of about 70 million would be awkward for the hypothesis — neither fish nor fowl. Even so, Hall *et al.*¹ boldly state that their discovery of a 15–25-million-year-old riodinid "provides additional support for a Gondwanan origin for many of the butterfly tribes and subfamilies". Is this really the case? And is there anything to suggest that butterflies are significantly older than 70 million years?

Within the true butterflies, the Papilionidae are the oldest group to diverge (Fig. 2). New molecular data seem to suggest that by far the largest genus in that group, *Papilio*, diverged about 55–65 million years ago¹⁰. Here again we have to beware of story-telling. Although this

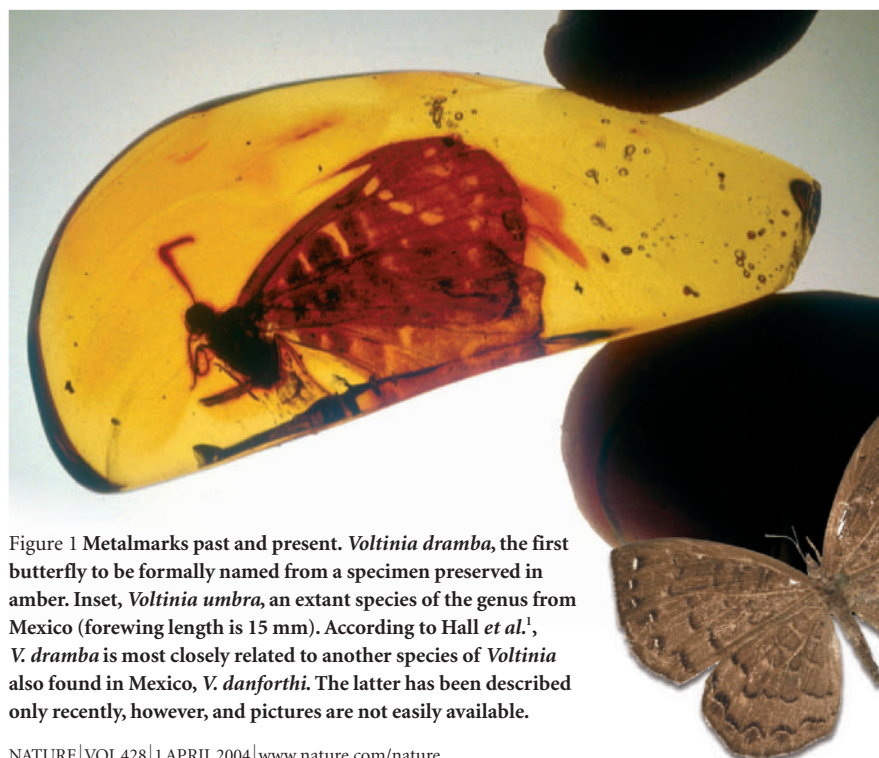


Figure 1 Metalmarks past and present. *Voltinia damba*, the first butterfly to be formally named from a specimen preserved in amber. Inset, *Voltinia umbra*, an extant species of the genus from Mexico (forewing length is 15 mm). According to Hall *et al.*¹, *V. damba* is most closely related to another species of *Voltinia* also found in Mexico, *V. danforthi*. The latter has been described only recently, however, and pictures are not easily available.

figure fits well with the 70-million-year approximation, divergence times based on molecular data are usually presented as 'error-less numbers' — with proper statistical caveats added, and considered in a geological context, such estimates often appear meaningless¹¹.

Whether or not molecular data will ever give reliable dates is a moot point, but I see no existing evidence to suggest that butterflies are older than about 70 million years, and little to imply a key role for Gondwana in their diversification. Most of the higher groups are either very widespread, or restricted to a single biogeographic region or continent¹². As de Jong has wryly observed⁹: "We have no idea when the butterflies originated, although there is no shortage of wild guesses." Nonetheless, by applying their outstanding knowledge of riodinid systematics to the description of this remarkable find, Hall *et al.*¹ bring new life to the study of fossil butterflies. But their work also reminds us of many issues, both methodological and philosophical, that continue to dog the purely observational sciences. ■

Dick Vane-Wright is in the Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.
e-mail: rivw@nhm.ac.uk

1. Hall, J. P. W., Robbins, R. K. & Harvey, D. J. *Proc. R. Soc. Lond. B* doi:10.1098/rspb.2004.2691 (2004).
2. Hedges, S. B. *Annu. Rev. Ecol. Syst.* **27**, 163–196 (1996).
3. May, R. M., Lawton, J. H. & Stork, N. E. in *Extinction Rates* (eds Lawton, J. H. & May, R. M.) 1–24 (Oxford Univ. Press, 1995).
4. Ehrlich, P. R. & Raven, P. H. *Evolution* **18**, 586–608 (1965).
5. Hall, J. P. & Harvey, D. J. *Mem. Entomol. Soc. Wash.* (in the press).
6. Owen, D. F. *Tropical Butterflies* (Oxford Univ. Press, 1971).
7. Ross, A. J., Jarzembowski, E. A. & Brooks, S. J. in *Biotic Response to Global Change: The Last 145 million Years* (eds Culver, S. J. & Rawson, P. F.) 288–302 (Cambridge Univ. Press, 2000).
8. Kristensen, N. P. & Skalski, A. W. in *Handbook of Zoology Vol. 4* (ed. Kristensen, N. P.) 7–25 (de Gruyter, Berlin, 1999).
9. de Jong, R. *Invertebr. Syst.* **17**, 143–156 (2003).
10. Zakharov, E. V., Caterino, M. S. & Sperling, F. A. H. *Syst. Biol.* (in the press).
11. Graur, D. & Martin, W. *Trends Genet.* **20**, 80–86 (2004).
12. Ackery, P. R., de Jong, R. & Vane-Wright, R. I. in *Handbook of Zoology Vol. 4* (ed. Kristensen, N. P.) 263–300 (de Gruyter, Berlin, 1999).
13. Vane-Wright, R. I. in *Butterflies: Ecology and Evolution Taking Flight* (eds Boggs, C. L., Watt, W. B. & Ehrlich, P. R.) 477–513 (Univ. Chicago Press, 2003).

trees. The cycads are a further case: the 120 or so living species of palm-like seed plants are the meagre remnant of a much more diverse group that is at least 250 million years old.

There are many more examples, and in textbooks the extant members of such groups are invariably described as having remained virtually unchanged for several hundred million years. Even if it is not stated explicitly, the implication is that these poor plants were forced into an evolutionary cul-de-sac by more successful plants, or by becoming too specialized.

In this context, however, the ferns are a special case. The fossil record shows them to be an old group (one type, the leptosporangiate ferns, which contain the majority of extant ferns, is more than 250 million years old). But with more than 10,000 extant species they remain fairly numerous. It has been proposed that these species are not just an ancient remnant, but the consequence of a more recent expansion. Did the ferns diversify relatively recently? Or are they just dwindling more slowly than some other groups?

To resolve these questions, several methods and data have to be used in combination. Schneider *et al.*¹ make good use of them. First, there must be a sufficiently detailed evolutionary tree — that is, a cladogram with branch lengths — depicting relationships and distances between the relevant groups. Second, a method for estimating the age of branches in the tree has to be available. Third, appropriate fossils are necessary to calibrate the tree, and make it a 'chronogram'; that is, a direct timescale is needed. Fourth, to have confidence in the age estimates, the analysis should indicate the margin of error in the estimates. At the extremes, combining all these requirements would be expected to show the extant lineages as old (long terminal branches; Fig. 2a) or much more recently diverged (short terminal branches; Fig. 2b).

Before evolutionary trees were in use, it was hard to establish even the order of events in evolution. When 'traits', such as the occurrence of mitochondria or flowers, were placed in their most optimal position on the evolutionary trees, it became possible to determine on which branches the traits evolved (and sometimes their order). But correlation of other kinds of events, and in particular correlations between lineages, are much more difficult. Such correlations usually involve time comparisons, which are problematic because evolutionary rates are commonly different among different lineages, and a general molecular clock — that is, one based on molecular changes and ticking at a constant rate — cannot be applied. So the goal of putting absolute times on the branches of the tree of life has been hampered by the lack of methods that use a variable molecular clock. Such methods do now exist^{2–4}, and the one used by Schneider *et al.* allows both for variation in evolutionary rates and inference of the level of

Evolutionary biology

Ferns reawakened

Torsten Eriksson

The principle of the evolutionary cul-de-sac is commonly invoked to explain the apparent lingering existence of once-diverse groups of organisms. Maybe that principle itself has had its day.

Some biological concepts keep popping up, even when they have been shown, time and again, not to be generally true. One well-known example is the 'biological species concept', the idea that only those organisms that can cross and produce fertile offspring belong to the same species. This can't generally be true for many reasons, the most obvious perhaps being that some organisms are not even sexual (such as bacteria and dandelions) and yet have species.

Schneider *et al.*¹ (page 553 of this issue) touch on another of these favourite concepts, the 'evolutionary cul-de-sac'. This is a common explanation for why some groups that show great diversity in the fossil record still exist but are greatly diminished in diversity, remaining largely unchanged — and supposedly unable to change. The new findings tell us that ferns, at least, do not belong in this category. Schneider *et al.* conclude that ferns (Fig. 1) have attained their current diversity much more recently than had been thought, and they probably did so as a response to the diversification of flowering plants.

During evolutionary history, many groups of organism have, of course, died out entirely. Plenty of others have persisted, however, even if much diminished compared with their apparent earlier diversity. Perhaps the best-known example among

land plants is the maidenhair tree (*Ginkgo*), which is the single living species of a lineage that is almost 300 million years old according to the fossil record. Horsetails (*Equisetum*) are another example: they now consist of only a handful of herbaceous species, but they belong to a lineage that was very diverse during the Carboniferous era (300 million years ago and older) and that included large



Figure 1 Ferns — diversified later than had been thought.

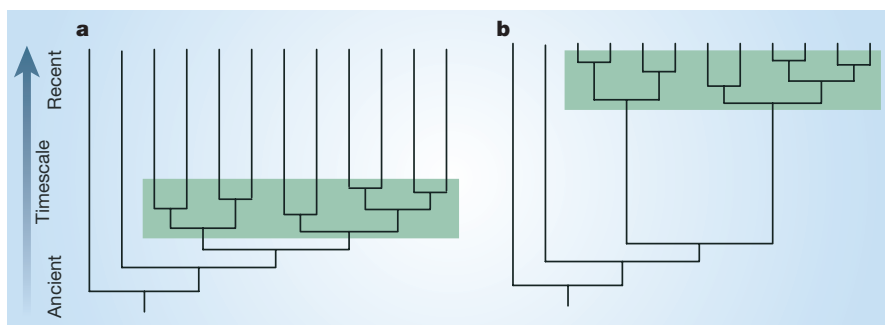


Figure 2 The long and the short of it. Two extreme outcomes following the time calibration of an evolutionary tree for a hypothetical group of organisms (shown in green). Diversification of the group was earlier in chronogram a than in chronogram b. As far as ferns are concerned, the results of Schneider *et al.*¹ indicate that the extant fern diversification pattern is more like that in b, whereas the traditional expectation would be that in a.

variability to accept from the data⁴. This may be a good compromise between enforcing a general molecular clock (which is still most commonly done) and allowing random variation of evolutionary rates.

What about the progress in analysing evolutionary trees? Until recently, it has been computationally almost impossible to reliably analyse the relationships of large numbers of species using model-based methods, such as 'maximum likelihood', which aims to find the statistically most likely tree. Model-based estimates of the amount of evolution on the trees (branch lengths) seem preferable, and new software with the 'bayesian inference of phylogeny' — as applied by Schneider *et al.* — is becoming increasingly used^{5,6}. This is partly because such methods can deal with the really large data sets that are

now being assembled, and still give information on how reliable the trees are. But there are also other reasons^{7,8}. One is that a bayesian-inference analysis yields not only a tree, but a sample of trees and model parameters, and this sample can be analysed to investigate various aspects of evolution. Schneider *et al.* used this sample in a novel way for the crucial purpose of obtaining standard deviations for their age estimates.

So, what about the results? Schneider *et al.* have added data on ferns to the huge database of sequence information on plant DNA that has built up during the past two decades. In their initial analysis, they dated the diversification of extant ferns as later than expected — around 80 million years ago, rather than perhaps three to four times that age. But they also went further to try to explain what made

that relatively recent diversification possible. By performing the same analysis on the flowering plants (angiosperms), they were able to correlate the age of the evolutionary 'radiations' of ferns and flowering plants, and they conclude that the extant ferns diversified after the angiosperms. The chronograms they produced appear on page 556.

Schneider and colleagues' explanation for fern diversification adds credibility to their conclusion: they argue that the expansion of flowering plants, and in particular the forests they created, increased environmental complexity and thereby the variety of habitats that could be explored by opportunistic organisms. Among such organisms were ferns, which in consequence underwent an evolutionary reawakening. It seems reasonable to assume that this could happen in any group when the opportunity arises, and it has indeed been shown to occur in certain tropical clubmosses (lycophods) that grow on trees⁹. Perhaps the whole idea of the evolutionary cul-de-sac is basically flawed. ■

Torsten Eriksson is at the Bergius Foundation, Royal Swedish Academy of Sciences, Box 50017, SE-10405 Stockholm, Sweden.

e-mail: torsten.eriksson@bergianska.se

1. Schneider, H. *et al.* *Nature* **428**, 553–557 (2004).
2. Thorne, J. L., Kishino, H. & Painter, I. S. *Mol. Biol. Evol.* **1**, 1647–1657 (1998).
3. Sanderson, M. J. *Mol. Biol. Evol.* **12**, 1218–1231 (1997).
4. Sanderson, M. J. *Mol. Biol. Evol.* **19**, 101–109 (2002).
5. Rannala, B. & Yang, Z. *J. Molec. Evol.* **43**, 304–311 (1996).
6. Huelsenbeck, J. P. & Ronquist, F. *Bioinformatics* **17**, 754–755 (2001).
7. Lewis, P. O. *Trends Ecol. Evol.* **16**, 30–37 (2001).
8. Huelsenbeck, J. P., Rannala, B. & Masly, J. P. *Science* **288**, 2349–2350 (2000).
9. Wikström, N. & Kenrick, P. *Mol. Phylogeny. Evol.* **19**, 177–186 (2001).

Nanoscale physics

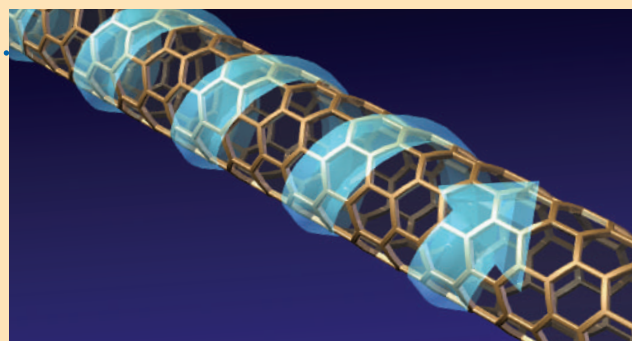
Big moment for nanotubes

As an electron whizzes around the nucleus of an atom, it develops a magnetic signal known as an orbital magnetic moment. The size of the moment depends on the outer diameter of the electron's orbit, which is fixed by the size of the atom. Electrons orbiting around the walls of a carbon nanotube (right) should also have an orbital magnetic moment, but this had previously never been detected. In this week's issue, E. D. Minot *et al.* are at last able to show that the resulting magnetic signal is exactly what's predicted (*Nature* **428**, 536–539; 2004).

Magnetic signals have been picked up in carbon nanotubes before, but because nanotubes come in several flavours (from metallic to semiconducting) and can have several layers (like a rolled cigar), it was not clear where the

magnetic moment was coming from. Minot *et al.* address this problem by using individual 'quasi-metallic' carbon nanotubes — which means that, although the nanotubes are not strictly metallic, it still takes only a small amount of energy to excite an electron so that it becomes a moving conduction electron.

Electrons on the threshold of freedom can orbit around the nanotube in a clockwise or anticlockwise direction; the orbital magnetic moments have the same magnitude for both directions, although their orientations are opposite. But when a magnetic field is applied, the electron energies are shifted. The clockwise orbits now have a different energy from the anticlockwise ones: half the electrons are shifted closer to the



conducting 'edge', while the others become harder to kick free.

The value of the energy shift depends on the strength of the applied field and on the orbital magnetic moment of the electrons, and Minot *et al.* were able to determine this precisely by measuring the conductance of a single nanotube suspended between two electrodes. The orbital magnetic moment derived for an electron traversing the waistline of a nanotube is about 10 to 20 times larger than that

for an electron in an atom — consistent with the difference in their diameters.

Minot and colleagues' conductance measurements are clear proof that orbital magnetic moments influence a nanotube's electron-energy levels in a magnetic field. They also raise the possibility of controlling the energy levels through an external magnetic field, opening the door to further studies of the fundamental properties of nanotubes, as well as technological applications.

May Chiao

Obituary

J. Beverley Oke (1928–2004)

Optical telescopes have very long useful lives, but they are only as good as the light-detecting instruments at their focal planes — which, in turn, are only as good as their designers. J. Beverley (Bev) Oke, who died on 2 March 2004 at the age of 75, devoted most of his scientific career to devising and building a succession of such instruments — photometers, cameras and spectrographs that have helped, in particular, to keep the 50-year-old Hale telescope at the Palomar Observatory, California, at the forefront of astronomical research.

Oke was not solely an instrument-builder; he was also interested in the data gleaned from the observations made with his instruments. His doctoral thesis, submitted at Princeton in 1953, was in fact a theoretical study, of the atmospheres of O-type stars — the most massive, hottest stars that burn hydrogen as their main energy source. Later, he moved into observational spectroscopy and this dominated his scientific interests for the rest of his career.

When Oke joined the Caltech faculty in 1958, he built a single-channel scanner, which could measure 10-nm-wide segments of the spectrum of a star or galaxy successively, by stepping a diffraction grating through the appropriate angle. Using this photometer, Oke became a key player in one of the most momentous events in modern astronomical history — the discovery, in 1963, of the large redshift of the quasar 3C273. (Redshift is a measure of an object's distance from the Earth.) Although we now know quasars to be extremely luminous, distant stars marking the centres of galaxies, at that time they were an enigma. Using the Hale telescope, Oke's Caltech colleague Maarten Schmidt identified the pattern of emission lines in the spectrum of quasar 3C273 as the Balmer series of lines from hydrogen atoms. The lines were difficult to recognize because they had been shifted to such long wavelengths — a consequence of the quasar's large redshift. Schmidt identified the Balmer series up to the so-called H β line from his photographic spectra of 3C273. Oke used his own photoelectric scanner to detect another line, H α ; his paper followed Schmidt's in *Nature*.

Oke went on to build a 'multi-channel scanner', a linear array of 33 photomultipliers, for the 200-inch Hale telescope. With colleagues, he used this device to make accurate measurements of the spectra of stars, Seyfert galaxies (tiny,



Astronomer and master instrument-builder

bright nuclei found in a few per cent of galaxies), quasars and the brightest galaxies in rich clusters over a range of redshifts. This last work, with James E. Gunn, was an attempt to determine the rate at which the expansion of the Universe is changing — brave though it was, it did not succeed. The changing expansion rate could not be measured because the luminosity of the brightest cluster galaxies changes with time or redshift. Only in the past decade has this classic cosmological measurement been made, using supernovae as 'standard candles'.

Oke always strove to make his observations as quantitative as possible: to do so meant facing the problem of calibrating observations to correct for the constantly changing, distorting effects of the Earth's atmosphere. The solution is to calibrate against the absolute fluxes of a few 'standard' stars, as they would be observed above the atmosphere. In a classic piece of work in 1970, Oke took his prime-focus photoelectric scanner from the 200-inch telescope and attached it to a 4-inch telescope at Palomar. His intention was to compare the flux from Vega, one of the brightest stars in the sky, at different wavelengths with the flux from a black-body cavity, held at the temperature

of molten platinum and mounted at the top of a pole at the observatory. By observing Vega at different 'air masses' as the atmosphere changed, and comparing it with the black body, Oke and Rudy Schild were able to eliminate the effects of the atmosphere and measure the absolute flux of Vega in physical units. Larger telescopes have since been able to make absolute-flux measurements of fainter stars and galaxies by observing both the source and the standard, Vega, through the atmosphere.

In the late 1970s, the charge-coupled device, or CCD, came into use in astronomy. The CCD offered a light-collection efficiency vastly better than that of the photographic plate, which had been the workhorse detector in astronomical applications for many decades. Oke built a revolutionary double spectrograph for Palomar, which split the light entering it into two beams. The beams were directed to two different CCDs, one optimized for blue light and one for red. This spectrograph is still in use today, although the CCD technology has been upgraded over the years. At the Keck Observatory in Hawaii, Oke was the principal investigator, with Judith Cohen, of the low-resolution imaging spectrograph (LRIS), mounted on one of the twin 10-m telescopes. LRIS contributed to many of the Keck telescopes' early successes, including the discovery and analysis of hundreds of galaxies at astonishingly high redshifts.

Oke retired from Caltech in 1992 to return to his native Canada. Until his death, he worked at the Dominion Astrophysical Observatory in Victoria, British Columbia, continuing his research into clusters of galaxies and rapid spectroscopy of cataclysmic variable stars. His scientific output was remarkable, given his commitment to teaching and to developing instrumentation.

Bev was a modest, phlegmatic man with a laconic sense of humour. His lectures were clear, matter of fact, and unadorned with fanciful speech. He led by example, not by fine words. His greatest indulgence was his MG sports car, which he delighted in driving along the S6 'highway to the stars' from Pasadena to Palomar Mountain. Until the very end, as well as fixing the MG, he was working on the design of a spectrograph for the proposed 30-m telescope, a US–Canadian project intended to achieve a resolution ten times better than that of the Hubble Space Telescope. It will be a fitting tribute.

Wallace Sargent

Wallace Sargent is in the Department of Astronomy, California Institute of Technology, Pasadena, California 91125, USA.
e-mail: www@astro.caltech.edu

Imaging

Mini X-ray tube sees the details

Appl. Phys. Lett. **84**, 2208–2210 (2004)

A miniaturized X-ray source developed by A. Haga *et al.* can produce X-ray images at very high resolution. The instrument reveals features as small as 20 μm across in microelectronic circuitry, making it potentially valuable for non-destructive evaluation of electronics. It also shows delicate bone structure and other fine details in biomedical imaging.

The X-rays are produced when a very fine beam of electrons hits a metal (tungsten) target. The key to the device is the narrow electron beam, which comes from carbon nanotubes on the conical tip of a palladium wire. The wire is coated with aligned nanotubes grown by chemical-vapour deposition of carbon. Each nanotube sprouts from a catalytic particle of palladium, which stays at the free end of the tube. These crystallites emit electrons, and strike the target 2 mm away; a small, evacuated alloy tube houses both components. Some of the X-rays generated escape through a window in the side of the tube to irradiate the sample.

Philip Ball

Ancient DNA

Etruscan origins

Am. J. Hum. Genet. **74**, 694–704 (2004)

The origins of the Etruscans, a civilization that came to dominate northern Italy from the eighth century BC, are a matter of much controversy. Cristiano Vernesi *et al.* have shed some light on the mystery by analysing ancient Etruscan DNA.

The researchers obtained mitochondrial DNA sequences from 80 Etruscans who lived between the seventh and third centuries BC, and whose skeletons are preserved in museum and public collections. They excluded samples that might have been

significantly degraded or contaminated with modern DNA.

When the sequences of individual Etruscans buried in different regions were compared, they showed relatively little genetic variation. This suggests that the Etruscans were a single population descended from common ancestors, rather than a collection of genetically different groups who shared a common culture.

When compared with sequences of modern-day Europeans, the Etruscans appear most closely related to populations on the southern and eastern Mediterranean shores. The best matches are with sequences from Tuscany and Turkey, hinting that the Etruscans evolved from local populations and that there was some interbreeding with populations in Asia Minor, which fits historical trading records. Nonetheless, few of the Etruscan sequences exactly match modern ones — signifying, perhaps, that the population rapidly died out after being swallowed by the Roman Empire. Helen Pearson

Plant biology

Softer pod shatters not

Cell **116**, 843–853 (2004)

Many plants depend on the wind, animals, birds or other agents to distribute their seeds. Others, such as the thale cress *Arabidopsis thaliana*, use mechanical means. Sarah J. Liljegren *et al.* have now identified an *Arabidopsis* gene, *INDEHISCENT*, that organizes this plant's spring-loaded dispersal mechanism.

Arabidopsis seeds, and those of its crop relatives, develop in a pod consisting of two elongated dish-shaped valves, separated by a thin cell layer, the replum. As they mature the valves dry and eventually fly open, throwing seeds some distance. The gene that Liljegren *et al.* have identified is required to form a narrow band of hardened tissue around the edges of the valves, as well as to weaken the join between valve and replum. This band shrinks less than the rest of the pod as it dries, creating the forces that shatter the pod.

Understanding how *INDEHISCENT* fits within the gene network that controls pod development will illuminate how plants form their elegant reproductive organs, and may also help farmers retain the 20–50% of oil-seed crops that, each year, are spilled on the ground.

Christopher Surridge

Cell biology

Ageing and actin

J. Cell Biol. **164**, 803–809 (2004)

Actin, a ubiquitous protein in eukaryotic organisms, forms long, dynamic filaments that are part of a cell's internal skeleton. Working with yeast, Campbell W. Gourelay *et al.* show that the stability of actin affects

the cell's lifespan. It apparently does so by modifying the release of reactive oxygen species from mitochondria — energy-producing bodies that make such species as a by-product.

Gourelay *et al.* grew mutant yeast that have a less dynamic actin skeleton than normal. Mitochondria were visualized with green fluorescent protein, and mitochondrial well-being was assessed by staining and imaging the cells. The mutants showed loss of mitochondrial function and increased levels of hydrogen peroxide, a common reactive oxygen species, and they died prematurely — apparently from the effects of the hydrogen peroxide.

In contrast, when the dynamic state of the actin was enhanced, as was seen by using a different actin mutant or deleting a gene encoding an actin stabilizer, the number of reactive oxygen species decreased and cell lifespan rose by up to 65%. Gourelay *et al.* point out that there is circumstantial evidence that the stability of the actin skeleton may influence cell death and lifespan in mammalian cells as well as in yeast.

Laura Nelson

Development

Cold comfort for chicks

J. Exp. Biol. **207**, 1543–1552; 1553–1561 (2004)

Cocooned under their warm, brooding mother, chick eggs do not need to regulate their own temperature, and are effectively cold-blooded. But when they hatch, chicks must regulate their own metabolism right from the start. It turns out that this ability develops during the chick's final days as an embryo, but only if the egg has had a cosy upbringing.

Juli L. Black and Warren W. Burggren incubated eggs at the usual temperature of 38 °C, and at 35 °C. They then measured the metabolic rates of late-stage embryos at temperatures down to 30 °C. The chicks raised at 38 °C maintained their oxygen consumption at around 30 microlitres per gram per minute throughout the late stage. But chicks incubated at 35 °C were unable to do the same. Their metabolic rates, and thus their heat production, declined as the ambient temperature dipped.

Black and Burggren next took tiny blood samples, at various developmental stages, from chick embryos incubated at the two temperatures. There was little difference between the two groups for much of gestation. But as the chicks entered the final stages of development, those incubated at 38 °C had significantly higher red-blood-cell counts, blood haemoglobin concentrations and oxygen-binding abilities than their cooler cousins, showing that a chilly incubation impairs blood development — and with it, the ability to thermoregulate through elevated metabolism.

Michael Hopkin



Etruscan art — detail from the Tomb of the Leopards, sixth century BC.

Ancient microRNA target sequences in plants

A gene-regulation mechanism in plants predates the emergence of flowering species.

MicroRNAs are an abundant class of small RNAs that are thought to regulate the expression of protein-coding genes in plants and animals. Here we show that the target sequence of two microRNAs, known to regulate genes in the class-III homeodomain-leucine zipper (HD-Zip) gene family of the flowering plant *Arabidopsis*, is conserved in homologous sequences from all lineages of land plants, including bryophytes, lycopods, ferns and seed plants. We also find that the messenger RNAs from these genes are cleaved within the same microRNA-binding site in representatives of each land-plant group, as they are in *Arabidopsis*. Our results indicate not only that microRNAs mediate gene regulation in non-flowering as well as flowering plants, but also that the regulation of this class of plant genes dates back more than 400 million years.

In angiosperms, class-III HD-Zip genes are required for meristem formation, establishment of adaxial domains during leaf development, and proper patterning and differentiation of vascular bundles^{1–5}. Dominant gain-of-function mutations in three *Arabidopsis* genes (*PHABULOSA*, *PHAVOLUTA* and *REVOLUTA*) map to a small region of the START domain, a domain that has sequence similarity to sterol-binding proteins in metazoans, suggesting that altered ligand binding might be responsible for the mutant phenotypes. But the discovery of two microRNAs with almost perfect complementarity to the messenger RNAs in the region of the gain-of-function mutations indicated that these phenotypes could arise from disruption of microRNA binding^{6,7}.

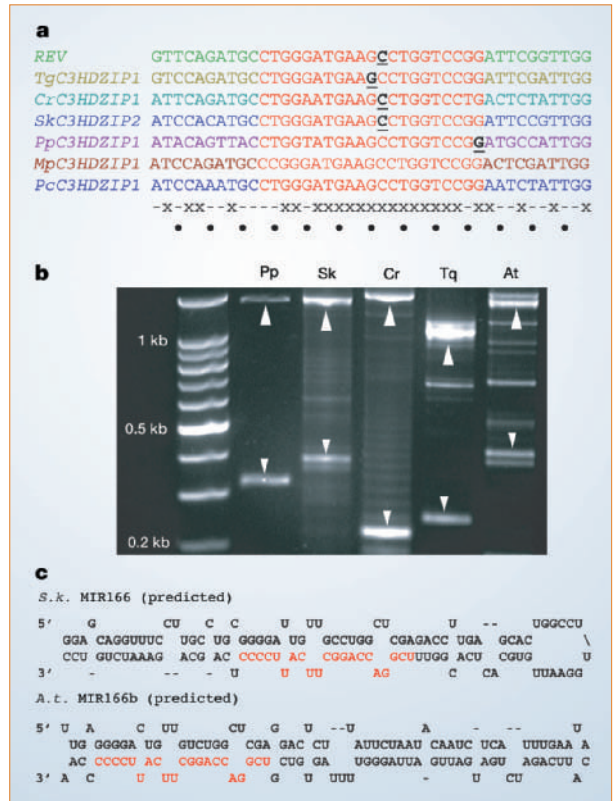
These microRNAs may negatively regulate target genes, perhaps through microRNA-directed cleavage within the region of complementarity^{8,9} or by disruption of translation¹⁰. *PHAVOLUTA* mRNA is cleaved *in vitro* within the predicted microRNA-binding site when incubated in a wheat-germ cell-free extract, and this cleavage is abolished by mutations conferring the dominant gain-of-function *PHAVOLUTA* phenotype⁸. Disruption of microRNA binding by altering the third base of a codon does not cause a change in the expressed amino-acid sequence and is sufficient to confer a dominant gain-of-function phenotype⁵.

Investigation of the phylogeny of this gene family in land plants revealed that the nucleotide sequence of the microRNA-binding site is conserved in class-III HD-Zip genes from two gymnosperms (*Pseudotsuga menziesii* and *Taxus globosa*), a fern (*Ceratopteris richardii*), a lycopod (*Selaginella kraussiana*), a moss (*Physcomitrella patens*), a liverwort

Figure 1 MicroRNA-mediated cleavage of class-III HD-Zip genes in land plants.

a, Alignment of sequences within the START domain of class-III HD-Zip genes from angiosperms (dark green), gymnosperms (light green), ferns (cyan), lycopods (purple), mosses (mauve), liverworts (brown) and hornworts (indigo). (See supplementary information for species' full names.) The region complementary to the *Arabidopsis* miR165/166 microRNA is in red. Crosses indicate identity in all land-plant genes included in our survey. Dots indicate third bases in codons. The 5' nucleotide of each predominant cleavage product detected in 5' RACE is underlined.

b, 5' RACE of mRNA from the indicated species. Bands indicated by large arrowheads represent full-length sequences; shortened fragments (small arrowheads) result from cleavage within the microRNA-binding site at positions shown in **a**. **c**, Predicted stem-loop structure containing the microRNA and its complement from the *Selaginella kraussiana* MIR166 homologue precursor complementary DNA sequence, and comparison with the predicted *Arabidopsis* MIR166b precursor⁷. The microRNA sequence is in red.



(*Marchantia polymorpha*) and a hornwort (*Phaeoceros carolinianus*). Even the base at the third position in codons is conserved, which would not normally be expected over such large timescales (Fig. 1a, and see supplementary information); within the binding site, 57% of codon third bases are identical.

Although a high degree of amino-acid conservation among these genes is found outside the microRNA-binding site, the nucleotide conservation is much lower — particularly at the third position in a codon. In the 57 conserved amino acids in the START domain (outside the microRNA-binding site), only 7.8% of the codon third positions are identical over all 16 sequences studied.

When cloning full-length complementary DNAs for these genes, we noted that shortened fragments were generated in 5' RACE (for rapid amplification of complementary DNA ends) experiments (Fig. 1b). Sequencing revealed that the longer products represent the 5' end of the full-length mRNA and that the shorter fragments are 3'-end cleavage products generated by microRNA-directed cleavage of the respective mRNAs^{8,9}. The cleavage site, which is in the middle of the microRNA-binding site (Fig. 1a, and see supplementary information), is nearly identical in all vascular plant taxa tested and is the same as that identified

by *in vitro* cleavage of *PHAVOLUTA* mRNA⁸.

We sequenced two separate products each for the *Selaginella* gene and *REVOLUTA*, and found that all four mapped to the same 5' end. By contrast, a putative cleavage site in the nonvascular plant *Physcomitrella* was discovered at the 3' end of the putative microRNA-binding site (Fig. 1a); five of six products sequenced had the site in this position. As the identity between the three *Physcomitrella* class-III HD-Zip sequences extends for at least 13 nucleotides from the 3' end to the conserved putative microRNA-binding site (see supplementary information), the difference in the location of the cleavage site could be due to a shift in the microRNA-binding site.

On the basis of an earlier observation that microRNA precursors are polyadenylated¹⁰, we cloned a full-length complementary DNA for MIR166 from *Selaginella kraussiana* by using primers derived from the sequence of *Arabidopsis* MIR165. Analysis of potential secondary structure using the web-based program *mfold*¹¹ revealed a predicted stem-loop structure involving the microRNA and a complementary region (Fig. 1c). Other than the MIR166 sequence and its hairpin complement, we detected no similarity between the *Selaginella* and predicted *Arabidopsis* MIR166 precursors.

Our findings indicate that all class-III

HD-Zip genes in land plants are negatively regulated through a conserved microRNA-mediated mechanism that has operated at least since the last common ancestor of bryophytes and seed plants, estimated to have existed more than 400 million years ago. The age of this mechanism is comparable to that of microRNA-mediated regulation of the *let-7* gene in metazoans¹².

Sandra K. Floyd, John L. Bowman

Section of Plant Biology, University of California,
Davis, California 95616, USA

e-mail: jlb Bowman@ucdavis.edu

1. Talbert, P. B., Adler, H. T., Parks, D. W. & Comai, L. *Development* **121**, 2723–2735 (1995).
2. McConnell, J. R. & Barton, M. K. *Development* **125**, 2935–2942 (1998).
3. McConnell, J. R. *et al. Nature* **411**, 709–713 (2001).
4. Baima, S. *et al. Development* **121**, 4171–4182 (1995).
5. Emery, J. F. *et al. Curr. Biol.* **13**, 1768–1774 (2003).
6. Rhoades, M. W. *et al. Cell* **110**, 513–520 (2002).
7. Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. & Bartel, D. P. *Genes Dev.* **16**, 1616–1626 (2002).
8. Tang, G., Reinhart, B. J., Bartel, D. P. & Zamore, P. D. *Genes Dev.* **17**, 49–63 (2002).
9. Llave, C., Xie, Z., Kasschau, K. D. & Carrington, J. C. *Science* **297**, 2053–2056 (2002).
10. Aukerman, M. J. & Sakai, H. *Plant Cell* **15**, 2730–2741 (2003).
11. Zuker, M. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
12. Pasquinelli, A. E. *et al. Nature* **408**, 86–89 (2000).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Medical genetics

A marker for Stevens–Johnson syndrome

Stevens–Johnson syndrome and the related disease toxic epidermal necrolysis are life-threatening reactions of the skin to particular types of medication^{1–3}. Here we show that there is a strong association in Han Chinese between a genetic marker, the human leukocyte antigen *HLA-B*1502*, and Stevens–Johnson syndrome induced by carbamazepine, a drug commonly prescribed for the treatment of seizures. It should be possible to exploit this association in a highly reliable test to predict severe adverse reaction, as well as for investigation of the pathogenesis of Stevens–Johnson syndrome.

We studied 44 patients with carbamazepine-induced Stevens–Johnson syndrome (CBZ–SJS), including five with overlapping toxic epidermal necrolysis, in whom the clinical morphology fulfilled Roujeau's diagnostic criteria¹ (see supplementary information). Patients suffered from widespread skin rashes with blisters, skin detachment and mucosa involvement. Controls included 101 patients who had been on carbamazepine for at least three months without adverse reaction ('CBZ-tolerant') and 93 normal individuals. All participants were Han Chinese residing in Taiwan. The study was approved by the institutional review board and informed consent was obtained in all cases.

Genetic factors influencing drug metab-

Table 1 Frequency of *HLA* alleles in patients with Stevens–Johnson syndrome

<i>HLA</i> allele	CBZ–SJS	CBZ-tolerant	Normal
<i>B*1502</i>	44 (100%)	3 (3%)*	8 (8.6%)*
<i>Cw*0801</i>	41 (93.2%)	17 (16.8%)	13 (14%)
<i>A*1101</i>	36 (81.8%)	51 (50.5%)	53 (57%)
<i>DRB1*1202</i>	33 (75%)	12 (11.9%)	18 (19.4%)
<i>B*1502, Cw*0801</i>	41 (93.2%)	3 (3%)	7 (7.5%)
<i>B*1502, A*1101</i>	36 (81.8%)	2 (2%)	6 (6.5%)
<i>B*1502, DRB1*1202</i>	33 (75%)	1 (1%)	5 (5.4%)
<i>B*1502, Cw*0801, A*1101, DRB1*1202</i>	29 (66%)	0 (0%)	3 (3.2%)

Frequencies (by number and percentage) of individual or combined loci of the *B*1502* ancestral haplotype are shown in patients with carbamazepine-induced Stevens–Johnson syndrome (CBZ–SJS; $n=44$), and in carbamazepine-tolerant ($n=101$) and normal subjects ($n=93$). For methods, see supplementary information.

*Odds ratio (CBZ–SJS/CBZ-tolerant): 2.504 (95% CI, 126–49,522); corrected P value $P_c=3.13 \times 10^{-27}$.

†Odds ratio (CBZ–SJS/normal): 895 (95% CI, 50–15,869); $P_c=1.38 \times 10^{-21}$.

olism and the immune response, including *HLA* genotype, might be involved in drug hypersensitivity^{4–6}. We therefore genotyped 157 cytochrome-P450 single-nucleotide polymorphisms using high-throughput MALDI–TOF mass spectrometry; we also genotyped all *HLA-B*, *-C*, *-A* and *-DRB1* alleles by sequence-specific oligonucleotide reverse line blot and sequence-based typing⁷.

We found that, compared with controls, there was no significant association between any of the cytochrome-P450 single-nucleotide polymorphisms and occurrence of CBZ–SJS. However, the alleles *B*1502*, *Cw*0801*, *A*1101* and *DRB1*1202* within the *HLA* region occurred at increased frequency in CBZ–SJS patients relative to the controls (Table 1). In particular, *HLA-B*1502* was present in 100% (44/44) of CBZ–SJS patients but in only 3% (3/101) of CBZ-tolerant patients and in 8.6% (8/93) of the general population.

When the CBZ-tolerant group is used as the control, the presence of *B*1502* has a 93.6% positive-prediction value for CBZ–SJS, whereas its absence has a negative-prediction value of 100%. In a test for CBZ–SJS, the *HLA-B*1502* allele should therefore have 100% sensitivity and 97% specificity.

After discovering six patients who were homozygous for the *HLA-B*1502* allele, we analysed the allele distribution in combined loci and discovered conserved alleles coupled at closely linked loci. This ancestral haplotype was defined as *B*1502, Cw*0801, A*1101, DRB1*1202*. It was present in 66% of the CBZ–SJS patients and in only 3% of the normal subjects, but was absent in CBZ-tolerant patients (Table 1).

The incidence of Stevens–Johnson syndrome in Han Chinese is higher than in Caucasians (8 cases per million person-years in Han Chinese compared with 2–3 cases in Caucasians). Carbamazepine is the drug most commonly associated with the syndrome in Asians (accounting for 25–33% of cases)^{8,9}, whereas only 5–6% of Caucasian CBZ–SJS cases are caused by it^{2,3}. We were unable to include Caucasian CBZ–SJS patients in our study; however, the presence of the *HLA-B*1502* allele in 8% of Han Chinese but only 1–2% of Caucasians¹⁰ may explain the lower incidence of CBZ–SJS

in Caucasians. Our findings could affect 1.2 billion Chinese worldwide; we do not yet know how they apply to other populations.

To our knowledge, the striking association of the allele *HLA-B*1502* with carbamazepine-induced Stevens–Johnson syndrome is the strongest so far described between an *HLA* marker and a disease; it even exceeds the classic example of *B27* and ankylosing spondylitis¹¹. It remains to be seen whether genes in the vicinity of the *HLA-B* locus, if not *B*1502* itself, participate in the pathogenesis of CBZ–SJS.

Wen-Hung Chung*, Shuen-Iu Hung†, Hong-Shang Hong*, Mo-Shong Hsih‡, Li-Cheng Yang*, Hsin-Chun Ho*, Jer-Yuarn Wu†§, Yuan-Tsong Chen†¶

Departments of *Dermatology and †Neurology, Chang Gung Memorial Hospital, Taipei, Taiwan
‡Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

e-mail: chen0010@ibms.sinica.edu.tw

§Department of Medical Research, China Medical University Hospital, Taichung, Taiwan

¶Department of Pediatrics, Duke University Medical Center, Durham, North Carolina 27710, USA

1. Roujeau, J.-C. *J. Invest. Dermatol.* **102** (suppl.), 28–30 (1994).
2. Roujeau, J.-C. *et al. N. Engl. J. Med.* **333**, 1600–1607 (1995).
3. Rzany, B. *et al. Lancet* **353**, 2190–2194 (1999).
4. Svensson, C. K., Cowen, E. W. & Gaspari, A. A. *Pharmacol. Rev.* **53**, 357–379 (2001).
5. Goldstein, D. B., Tate, S. K. & Sisodiya, S. M. *Nature Rev. Genet.* **4**, 937–947 (2003).
6. Roujeau, J.-C. *et al. Tissue Antigens* **28**, 251–254 (1986).
7. Erlich, H. A., Opelz, G. & Hansen, J. *Immunity* **14**, 347–356 (2001).
8. Khoo, A. K. M. & Foo, C. L. *Burns* **22**, 275–278 (1996).
9. Kamaliah, M. D., Zainal, D., Mokhtar, N. & Nazmi, N. *Int. J. Dermatol.* **37**, 520–523 (1998).
10. Geer, L., Terasaki, P. I. & Gjertson, D. W. in *HLA 1998* (eds Gjertson, D. W. & Terasaki, P. I.) 327–363 (Am. Soc. Histocompat. Immunogenet., Kansas, 1998).
11. Schlosstein, L., Terasaki, P. I., Bluestone, R. & Pearson, C. M. *N. Engl. J. Med.* **288**, 704–706 (1973).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

brieF communications arising online

www.nature.com/bca

Tumour suppression: Disruption of *HAUSP* gene stabilizes p53

Jordan M. Cummins, Carlo Rago, Manu Kohli, Kenneth W. Kinzler, Christoph Lengauer, Bert Vogelstein (doi:10.1038/nature02501)

Tumour suppression

Disruption of *HAUSP* gene stabilizes p53

Arising from: Li, M. *et al. Nature* **416**, 648–653 (2002).

Ubiquitination of p53 is the principal mechanism through which p53 concentrations in the cell are regulated^{1,2} in order to maintain its effects on tumorigenesis and normal cell growth. The protein HAUSP (also known as USP7) is a ubiquitin-specific protease (deubiquitinase)^{3,4} that has been shown by Li *et al.* to bind to p53 (ref. 5); in overexpression experiments, Li *et al.* showed that p53 could be stabilized as a result of deubiquitination by HAUSP and suggested that HAUSP may thereby act as a tumour suppressor^{5,6}. Here we use a different approach to investigate the relationship between HAUSP and p53 stability, in which we disrupt the *HAUSP* gene in human cells by targeted homologous recombination. Instead of the expected increase in ubiquitinated p53 and destabilization of p53, we find that disruption of *HAUSP* results in the opposite phenotype, leading to stabilization and functional activation of p53 in our system. It may be that HAUSP can deubiquitinate other proteins such as MDM2, another regulator of p53, and that the balance between the deubiquitination of the different targets of HAUSP determines the steady-state level of p53.

We sequentially targeted the two alleles of *HAUSP* in the human colorectal cancer cell line HCT116. This line was chosen because the *p53* gene and its cellular pathways have been shown to perform normally⁷ in it. The rate of generating homozygous 'knock-out' (*HAUSP*^{-/-}) clones was exactly that predicted from the rate of generating heterozygous (*HAUSP*^{+/-}) clones, providing cogent evidence against clonal selection bias. The absence of HAUSP protein in the knockout clones was confirmed by western blotting (Fig. 1a).

Surprisingly, the absence of HAUSP resulted in a substantial increase in the steady-state level of p53 protein (Fig. 1a); this result was confirmed by the strong immunofluorescence of *HAUSP*^{-/-} compared with *HAUSP*^{+/+} HCT116 cells obtained with a p53-specific antibody (DO-1) and an Alexa 488-conjugated anti-mouse secondary antibody (results not shown). The p53 protein was functional, as assessed by expression of p21, a gene that is transcriptionally activated by p53 (ref. 8; Fig. 1a). The p21 protein inhibits cell-cycle progression, causing *HAUSP*-disrupted cells to grow markedly more slowly than parental cells (Fig. 1b).

To test the generality of these results, we attempted to disrupt the *HAUSP* gene in two other lines with intact p53 function, one derived from a cancer cell line (SW48) and the other derived from telomerase-immortalized, normal retinal-pigment epithelial

cells (RPE). One allele of *HAUSP* could be efficiently disrupted in SW48 (four clones derived from 1.4×10^7 cells). After Cre-mediated excision of the sequences conferring geneticin resistance within the targeted allele, the *HAUSP*^{+/-} cells were used in a second round of targeting. Although a high efficiency of homologous recombination was seen (13 clones derived from 4.2×10^7 cells), all of the integrations proved to be in the allele that was disrupted in the first round of targeting.

In RPE cells, the first allele was also easily disrupted (two clones from 1.4×10^7 cells). After Cre-mediated excision of the geneticin-resistance gene, we obtained 64 clones with homologous integration at the *HAUSP* locus. Fifty-two of these had reintegrated in the original targeted allele

(and so remained *HAUSP*^{+/-}), whereas 12 had reintegrated in the opposite allele (and were therefore *HAUSP*^{-/-}). Although all *HAUSP*^{+/-} reintegrants were expandable, the 12 *HAUSP*^{-/-} clones would not proliferate after first passage. We were able to prepare proteins from only one of the *HAUSP*^{-/-} clones before its demise. Western blotting confirmed that no HAUSP expression was detectable and, as in *HAUSP*^{-/-} HCT116 cells, p53 levels were increased and p21 was accordingly upregulated (Fig. 1a).

We next investigated whether the up-regulation of p53 following disruption of *HAUSP* was associated with altered ubiquitination. Ubiquitinated forms of p53, which manifest themselves as a series of slowly migrating bands that react with a p53-specific antibody, were observed in parental HCT116 cells (Fig. 1c). This ubiquitination was decreased in *HAUSP*^{-/-} cells, the opposite of what would be expected if the principal function of HAUSP were to deubiquitinate p53. Proteasome

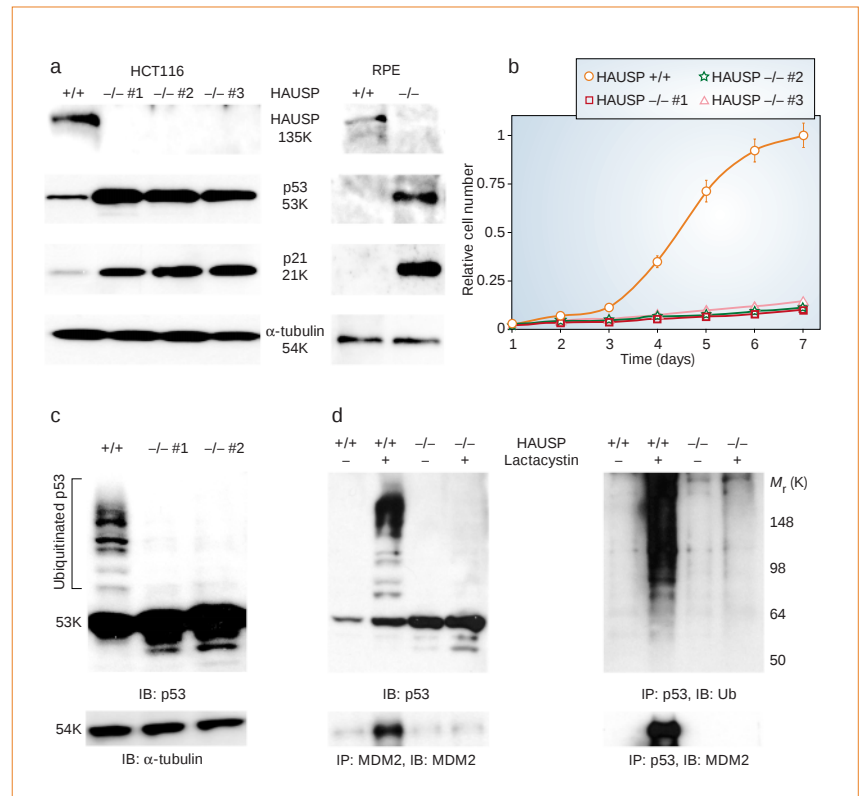


Figure 1 Destabilization of p53 by HAUSP. **a**, Immunoblots of three different clones of HCT116 and one clone of RPE *HAUSP*^{-/-} cells, using the antibodies indicated. Antibody against HAUSP was generously provided by R. Everett³. **b**, Growth assay depicting the slow proliferation of *HAUSP*^{-/-} cells. Growth was measured by assessing DNA content using SYBR Green I. The average fluorescence of 12 replicates of each cell type was measured (starting with about 3,000 cells per well on day 0). **c**, A p53 immunoblot (IB), when overexposed, shows slowly migrating forms of p53 resulting from ubiquitination in *HAUSP*^{+/+} but not *HAUSP*^{-/-} cells; α-tubulin was used as a loading control. **d**, Lysates from untreated cells and from cells treated with lactacystin were immunoblotted (IB) with an antibody against p53, or immunoprecipitated (IP) with antibodies against p53 or MDM2 and immunoblotted with a ubiquitin-specific antibody (Ub) or an antibody against MDM2. The strategy used for creating knockouts with AAV vectors¹¹ is described in ref. 12. The targeting construct pAAV-neo-HAUSP was made by polymerase chain reaction, using bacterial artificial chromosome clone RPC111.c77H9 (Invitrogen) as the template for the homology arms. A targeted insertion was made in the beginning of exon 6 (essential for deubiquitinase activity). Stable G418-resistant clones were initially selected in the presence of 0.4 mg ml⁻¹ Geneticin (Invitrogen), then routinely propagated in the absence of selective agents. Standard procedures were followed for cell culture and immunoblotting. Further details are available from the authors.

inhibitors such as lactacystin prevent the degradation of ubiquitinated proteins and therefore accentuate differences in ubiquitination. After treatment with lactacystin, a striking difference in the ubiquitination of p53 was observed between *HAUSP*^{+/+} and *HAUSP*^{-/-} cells; HAUSP seemed to stabilize MDM2, the E3 ubiquitin-ligase responsible for p53 destabilization (Fig. 1d).

Our results establish that HAUSP is required for proper p53 regulation, but in the opposite direction to that predicted by the seminal observations of Li *et al.*⁵ In the absence of HAUSP, p53 levels increased as its ubiquitination decreased. This resulted in slow growth (in the case of HCT116 cells) or failure to thrive (in SW48 and RPE cells), which are phenotypes reminiscent of those achieved by MDM2 targeting in mice^{9,10}. These results indicate that HAUSP may be able to deubiquitinate

other targets, particularly MDM2, and that the net deubiquitination of the various targets of HAUSP determines the steady-state level of p53.

Note added in proof: A related paper on the role of HAUSP in the p53–MDM2 pathway has appeared elsewhere (M. Li, C. Brooks, N. Kong & W. Gu *Mol. Cell* **13**, 879–886; 2004).

Jordan M. Cummins, Carlo Rago, Manu Kohli, Kenneth W. Kinzler, Christoph Lengauer, Bert Vogelstein

Howard Hughes Medical Institute, The Sidney Kimmel Comprehensive Cancer Center, and Program in Cellular and Molecular Medicine, The Johns Hopkins University Medical Institutions, Baltimore, Maryland 21231, USA
e-mail: vogelbe@welch.jhu.edu
doi:10.1038/nature02501

1. Michael, D. & Oren, M. *Semin. Cancer Biol.* **13**, 49–58 (2003).

2. Haupt, Y., Robles, A. I., Prives, C. & Rotter, V. *Oncogene* **21**,

8223–8231 (2002).

3. Everett, R. D. *et al. EMBO J.* **16**, 1519–1530 (1997).

4. Hu, M. *et al. Cell* **111**, 1041–1054 (2002).

5. Li, M. *et al. Nature* **416**, 648–653 (2002).

6. Wood, S.A. *Sci. STKE* 2002 PE34 (2002).

7. Bunz, F. *et al. J. Clin. Invest.* **104**, 263–269 (1999).

8. Vogelstein, B., Lane, D. & Levine, A. J. *Nature* **408**, 307–310 (2000).

9. Jones, S. N., Roe, A. E., Donehower, L. A. & Bradley, A. *Nature* **378**, 206–208 (1995).

10. Montes de Oca Luna, R., Wagner, D. S. & Lozano, G. *Nature* **378**, 203–206 (1995).

11. Hirata, R., Chamberlain, J., Dong, R. & Russell, D. W. *Nature Biotechnol.* **20**, 735–738 (2002).

12. Kohli, M., Rago, C., Lengauer, C., Kinzler, K.W. & Vogelstein, B. *Nucleic Acids Res.* **32**, E3 1–8 (2004).

Competing financial interests: declared. (Under a licensing agreement between Genzyme Corporation, CalBioChem and the Johns Hopkins University, K.W.K. and B.V. are entitled to a share of royalty received by the university on sales of p53-related and p21-related products described in this communication. The university and these investigators also own Genzyme Corporation stock, which is subject to certain restrictions under university policy. K.W.K. is a paid consultant to, and receives research funding from, Genzyme. The terms of this arrangement are being managed by the Johns Hopkins University in accordance with its conflict-of-interest policies.)

Designing materials for biology and medicine

Robert Langer¹ & David A. Tirrell²

¹Department of Chemical Engineering, Massachusetts Institute of Technology, Building E25-342, Cambridge, Massachusetts 02139, USA

²Division of Chemistry and Chemical Engineering, California Institute of Technology, Mail Code 164-30, Pasadena, California 91125, USA

Biomaterials have played an enormous role in the success of medical devices and drug delivery systems. We discuss here new challenges and directions in biomaterials research. These include synthetic replacements for biological tissues, designing materials for specific medical applications, and materials for new applications such as diagnostics and array technologies.

Biomaterials have been defined as substances other than foods or drugs contained in therapeutic or diagnostic systems¹ and, in some cases, have been described as materials composed of biologically derived components (for example, amino acids) irrespective of their application. Throughout history, biomaterials have played an important role in the treatment of disease and the improvement of health care. Early biomaterials include metals such as gold that were used in dentistry over 2,000 years ago. Other early examples of biomaterials include wooden teeth and glass eyes². However, with the advent of synthetic polymers at the end of the nineteenth century, these materials became increasingly used in health care. For example, polymethylmethacrylate, PMMA, was used in dentistry in the 1930s¹ and cellulose acetate was used in dialysis tubing² in the 1940s. Dacron was used to make vascular grafts; polyether-urethanes, the materials used in ladies' girdles, were used in artificial hearts; and PMMA and stainless steel were used in total hip replacements^{1,2}. Naturally occurring materials such as collagen have also been used as biomaterials³. However, in nearly every case, these materials were adopted from other areas of science and technology without substantial redesign for medical use. Although these materials helped usher in new medical treatments, critical problems in biocompatibility, mechanical properties, degradation and numerous other areas remain. To this end, scientists are creating new materials including those with improved biocompatibility, stealth properties, responsiveness (smart materials), specificity and other critical properties. Modern biomaterials science is characterized by a growing emphasis on identification of specific design parameters that are critical to performance, and by a growing appreciation of the need to integrate biomaterials design with new insights emerging from studies of cell–matrix interactions, cellular signalling processes, and developmental and systems biology.

Biomaterials are already having an enormous effect on medicine. Controlled drug delivery systems that largely involve polymers⁴ are used by tens of millions of people annually⁵. Recent examples are polymer-coated stents, which have recently been approved both in Europe and the United States. Hundreds of thousands of lives are expected to be saved each year⁶. In addition, various controlled release systems for proteins, such as human growth hormone, as well as molecules decorated with polyethylene glycol (PEG), such as pegylated interferon^{4–7}, have recently been approved by regulatory authorities, and are showing how biomaterials can be used to positively affect the safety, pharmacokinetics and duration of release of important new drugs. Another area where biomaterials have recently had an impact is in tissue engineering. By combining polymers with mammalian cells, it is now possible to make skin for patients who have burns or skin ulcers, and various other polymer/cell combinations are in clinical trials, including corneas, cartilage, bone and liver⁸. Biomaterials have also had a major impact as the central components of dental implants, sutures, and numerous medical devices².

Here we describe novel material concepts that are shaping future directions in biomaterials science. In particular, we discuss (1) creating synthetic replacements for biological tissues using naturally occurring building blocks, (2) synthesizing materials using man-made building blocks for specific medical and biological applications, and (3) design concepts for new *in vitro* applications such as diagnostics and array technologies.

Synthetic replacements for biological tissues

Materials composed of naturally occurring (biologically derived) building blocks, including extracellular matrix (ECM) components, are being studied for applications such as direct tissue replacement and tissue engineering. The ECM, a complex composite of proteins, glycoproteins and proteoglycans, provides an important model for biomaterials design⁹. ECM-derived macromolecules (for example, collagen) have been used for many years in biomaterials applications³, and it is now possible to create artificial analogues of ECM proteins using recombinant DNA technology¹⁰. Through the design and expression of artificial genes, it is possible to prepare artificial ECM proteins with controlled mechanical properties and with domains chosen to modulate cellular behaviour. This approach avoids several important limitations encountered in the use of natural ECM proteins, including batch-to-batch (or source-to-source) variation in materials isolated from tissues, restricted flexibility in the range of accessible materials properties, and concerns about disease transmission associated with materials isolated from mammalian sources.

Elastin-based systems have been of special interest in this regard. Urry and co-workers have shown over many years that simple repeating polypeptides related to elastin can be engineered to exhibit mechanical behaviour reminiscent of the intact protein¹¹. Crosslinking can be accomplished via radiative¹² or chemical¹³ means, and electrospinning has been used to prepare fibrous forms of engineered elastins¹⁴ (Fig. 1). Incorporation of cell-adhesion ligands allows attachment and spreading of cultured cells, and in the specific case of materials for vascular grafts, retention of endothelial cell adhesion in the face of shear stresses characteristic of the normal circulation¹⁵.

The promise of biosynthetic approaches to biomaterials design must be weighed against the fact that very little is known about the *in vivo* performance of systems prepared in this way. Animal experiments on elastin-like polypeptides prepared by chemical synthesis have shown these materials to evoke relatively mild tissue reactions, but the range of materials investigated to date has been small¹⁶. As more is learned about the mechanical properties of such materials and about their interactions with cells in culture, the groundwork will be laid for more extensive evaluation in animal models. Recent progress in the development of methods for incorporation of non-natural amino acids into recombinant proteins points the way to an alternative strategy for preparing artificial ECM proteins with diverse chemical, physical and biological properties¹⁷.

Substantially more experience has been gained in evaluating the *in vivo* performance of engineered biomaterials based on polysaccharides. Alginate hydrogels bearing cell-adhesion ligands have been used as scaffolds for cell encapsulation and transplantation, and have yielded promising results in experiments directed towards the engineering of bone tissue capable of growth from small numbers of implanted cells¹⁸. The prospect of growing tissues from small numbers of precursor cells is an attractive alternative to harvesting and encapsulating large cell masses before transplantation.

Molecular self assembly of peptides or peptide-amphiphiles may also lead to unique biomaterials. A number of self assembled peptide systems have been developed, including systems that can potentially be used in tissue engineering and nanotechnology^{19,20}.

An alternative to synthesizing polymers composed of natural components is the synthesis of biomimetic polymers, which combine the information content and multifunctional character of natural materials (such as a particular amino acid sequence that might be desirable for cell attachment) with the tailorability of a synthetic polymer, such as control of molecular mass or polymer degradation, and the ability to impart appropriate mechanical properties. An example of this concept has been the synthesis of polymers composed of lactic acid and lysine. Like polylactic-glycolic acid, these polymers can be made to degrade at desired times. However, by adding lysine as a co-monomer with lactic acid, a free amino acid is provided, which allows coupling reactions to take place and does not affect the overall biocompatibility of the polymer²¹. By coupling specific amino acids (such as the tripeptide sequence RGD) to this polymer, cell adhesion can be regulated²².

Another strategy that has been used in modifying a variety of natural and synthetic polymers has been the inclusion of PEG into the material to reduce non-specific effects of protein adsorption and colloidal aggregation. The molecular origins of these phenomena are not yet thoroughly understood. One example of the effects of the PEG can be seen in the creation of polylactic-glycolic acid (PLGA) PEG diblock polymers that, when formed into nanospheres, can, like cells, circulate in the body for long time periods. For example, in mice, only 30% of PLGA-PEG nanoparticles were cleared after 5 h whereas 66% of non-PEG-containing PLGA nanoparticles were

cleared in only 5 min (ref. 23). Halstenberg *et al.*²⁴ have adopted a related approach in engineering protein-based biomaterials, by grafting poly(ethylene glycol) diacrylate onto an artificial protein that contained multiple cysteine residues. The protein was designed to serve several functions, including cell adhesion, heparin binding, and degradation by plasmin to facilitate penetration by invading cells. The mechanical properties of the material were controlled by photopolymerization of the pendant acrylate units.

Another approach to creating a biomimetic reversible system is the creation of an antigen responsive hydrogel. Corresponding antibody pairs are used to form reversible non-covalent crosslinks in a polyacrylamide system. In the presence of excess free antigen, the hydrogel swells, but in its absence, the gel collapses back to a crosslinked network. Swelling does not occur when foreign antigens are added, showing that the system is antigen specific. Release of a model protein such as haemoglobin has been demonstrated in response to specific antigens²⁵.

Materials for specific medical and biological applications

A variety of new materials are being synthesized from man-made building blocks, and being used to create devices for specific medical applications. One area of increasing attention has been the development of shape-memory materials that have one shape at one temperature and another shape at a different temperature²⁶. Such materials might permit new medical procedures. For example, current approaches for implanting medical devices often require complex surgery followed by device implantation. However, with the development of minimally invasive surgery, it is possible to place small devices inside the body using laparoscopes. These types of surgical advances may create new opportunities to enable a bulky device to be implanted into the human body in a convenient way.

Shape-memory materials might provide such an opportunity, because they have the ability to memorize a permanent shape that can be substantially different from an initial temporary shape. Thus bulky devices could potentially be introduced into the body in a temporary shape, like a string, that could go through a small laparoscopic hole, but then be expanded on demand into a permanent shape (for example, a stent, a sheet, and so on) at body temperature (Fig. 2). New polymers have been synthesized with this concept in mind, including phase segregated multiblock copolymers whose starting materials are known biocompatible monomers such as ϵ -caprolactone and *p*-dioxanone. Generally, these materials have at least two separated phases, each with thermal transition (glass or melting) temperatures. The phase with the higher transition temperature is responsible for the permanent shape, whereas the second phase can act as a molecular switch and enable fixation of a temporary shape. By regulating temperatures above and below the second phase's transition temperature, shape can be shifted from one form to another. In addition to permitting new opportunities for implanting devices, these polymers have been developed into sutures that are able to tie themselves on demand, triggered by a temperature change²⁷ (Supplementary video 1).

Materials that are liquids at room temperature but that harden in response to a change in temperature or an external stimulus, such as light, are also being studied^{28–30}. Such systems offer an opportunity to inject materials containing drugs, for example, into the body through a small needle, but still enable the formation of an implant.

A variety of smart gels have also been developed. These gels can respond (and swell, for example) to triggers such as temperature or pH, or even specific molecules in the body such as glucose. Such systems may, with further study, be valuable in the treatment of disease: in the case of diabetes, for example, smart gels could provide direct feedback control, allowing more insulin to be delivered in response to excess glucose³¹.

It may be possible to change not only the bulk properties of materials but also their surface properties, using a simple 'switch'

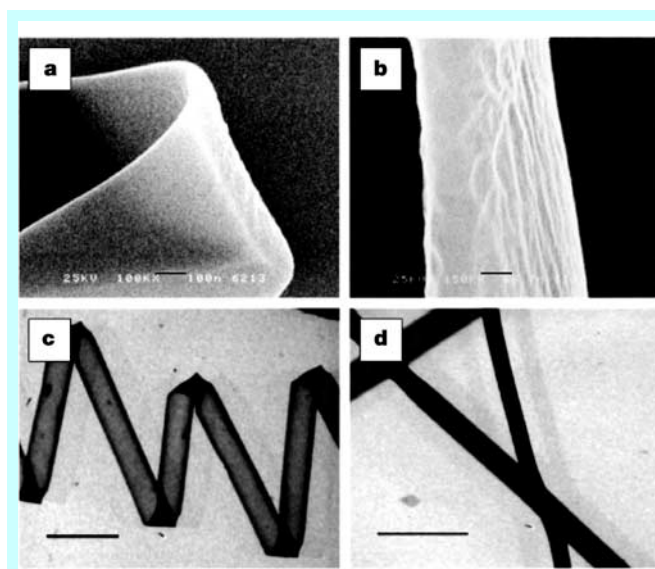


Figure 1 Electrosun fibres of elastin-like artificial proteins made by expression of artificial genes in bacterial cells. **a, b**, Scanning electron micrographs. **c, d**, Transmission electron micrographs. Scale bars: **a** and **b**, 100 nm; **c**, 3.3 μ m; **d**, 2.0 μ m. Images from ref. 14.

such as temperature or electric charge. For example, alkanethioates such as 16-mercaptohexadecanoic acid, which has a hydrophobic chain capped by a hydrophilic carboxyl group, forms self-assembled monolayers on gold surfaces. These chains have an upright equilibrium conformation presenting the carboxylic groups to the surrounding medium. But when these alkanethioates are placed on gold at the correct density, application of an electric potential causes the carboxyl groups to be attracted to the gold surface electrostatically, causing the molecules to reversibly rearrange and expose the hydrophobic chain (Supplementary video 2). Such surface switches might offer new opportunities in such areas as drug delivery, microfluidics and biosensors³².

The development of high-throughput approaches to create novel biopolymers and screen them for various applications is garnering increased attention. For example, Kohn and co-workers have created polymer libraries and then screened them for different applications^{33,34}. This type of high-throughput approach has also been used in the creation of gene therapy agents. For example, poly- β -amino esters can be synthesized in a high-throughput manner, and a number of these new polymers have been shown to have higher DNA transfection activities in cell-based assays than existing materials such as lipofectamine^{35,36}.

Gene therapy represents an area where appropriate molecular design is critical to achieving a successful outcome. Although viral vectors are highly effective, their use has raised serious safety concerns. This has motivated research on synthetic gene therapy vectors, which, although safer, have thus far been much less effective than viral vectors. To be effective, there are a number of attributes that the material must possess, including the ability to condense DNA to sizes less than 150 nm so that it can be taken up by receptor-mediated endocytosis, the ability to be taken up by endosomes in

the cell and to allow DNA to be released in active form, and to enable it to travel to the cell's nucleus³⁷. An interesting example of the design of such new materials is provided by the cationic-cyclodextrin polymers developed by Davis and co-workers. Cyclodextrins are relatively non-toxic and do not elicit an immune response. When Davis and co-workers initially used cyclodextrins for packaging DNA they found the resulting complexes to be relatively unstable. To address this, they added adamantane-conjugated polyethylene glycol to the surface of the cyclodextrin particles. This enabled the development of uniformly sized nanoparticles that resisted aggregation. By modifying the surface of the particles in this way, chemical groups could be exposed that could attach other molecules and allow the particles to be targeted to, and deliver genes to, specific cells³⁸. Another novel approach for gene therapy involves creating a triplex where low-density lipoprotein is used for targeting and stearyl polylysine is used for DNA complexation. This approach has been used to deliver vascular endothelial cell growth factor to heart muscle to aid in treating blockage of blood vessels³⁹. With the advent of new gene therapy agents such as RNA interference⁴⁰, the ability to design improved materials to deliver these agents will have increasing importance.

Biomaterials are also being used to affect bioadhesion in novel ways, even enabling materials to be used to potentially deliver complex molecules orally. It appears that polymers with high concentrations of hydroxyl groups bind with the intestinal mucosa. For example, Mathiowitz and co-workers⁴¹ have designed polyanhydride nanoparticles (which can encapsulate DNA or other molecules, and which expose carboxyl groups on their exterior as the polymer erodes), and have shown that they can attach to mucous membranes and bind to the intestinal wall. Poly(fumaric-co-sebacic) anhydride showed greater adhesive forces than other materials tested, had a longer contact time with cells in the intestine, and was able to pass through the intestinal wall with a protein inside it. Peppas and co-workers have developed polymers that are not only bioadhesive, but also swell in response to a pH change. These polymers are able to protect proteins from the acidic pH of the stomach, and release it in the more basic pH of the intestine. These materials also appear to temporarily open connections between intestinal cells, allowing the proteins to pass through⁴².

Microfabrication-based devices may also provide a novel approach for creating a variety of new biomaterials and delivery systems. For example, silicon microchips have been engineered to contain over 100 nanometre-sized drug-containing wells covered with gold on a chip 1 cm $\times$ 1 cm. By applying approximately 1 V to individually addressable wells, the drug in any of these wells can be released⁴³. These types of systems have also been recently made with degradable materials (Fig. 3)⁴⁴.

Microfabrication has also been used in the creation of sensors. For example, small sensors are being used to measure intraocular pressure for glaucoma patients⁴⁵, and silicone microstimulators have been developed that can be controlled by telemetry⁴⁶ and are being used for retinal stimulation to aid in photoreceptor generation to treat diseases of the back of the eye that cause blindness. Prausnitz and co-workers have developed microneedles that are able to penetrate into the skin to depths far enough for drugs to enter the circulation, but shallow enough not to reach skin layers that contain nerves: therefore they do not cause pain⁴⁷. Microfabrication has also been used to create polymer scaffolds containing an intricate vascular network⁴⁸. All of these represent important ways of using micro- and nanotechnology to create potential new medical devices.

New applications: diagnostics and array technologies

Much of the preceding discussion has focused on the development of materials that display biological information, often in the form of peptide or protein domains, for the purpose of controlling cell and tissue behaviour. A related challenge arises in the engineering of

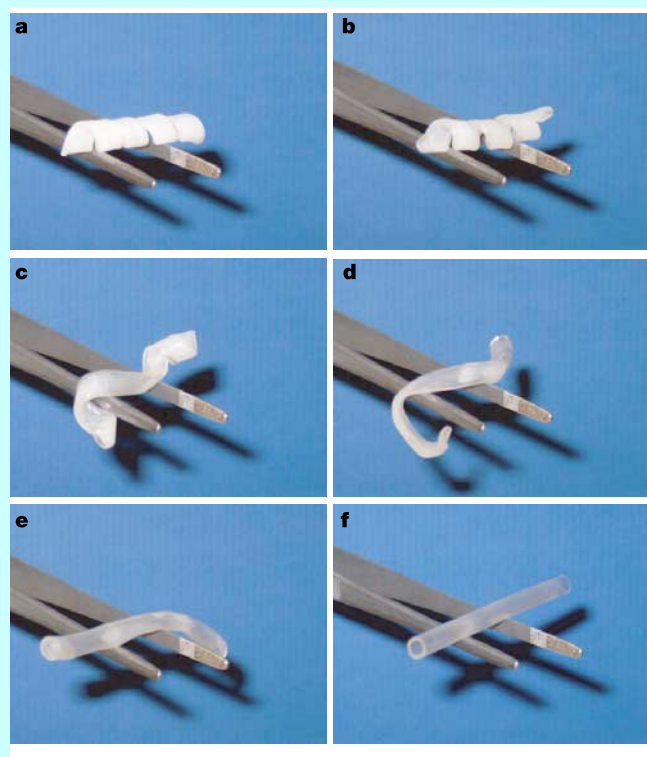


Figure 2 Time series of photographs showing recovery of a shape-memory tube. **a–f**, Start to finish of the process; total time, 10 s at 50 °C. The tube was made of a poly(ϵ -caprolactone)dimethacrylate polymer network (the M_n of the network's switching segments was 10^4) that had been programmed to form a flat helix. Images courtesy of A. Lendlein.

materials for diagnostics and array technologies, in which large numbers (typically hundreds or thousands) of nucleic acids⁴⁹ or proteins⁵⁰ are presented in a format that allows rapid and highly parallel read-out of information concerning gene expression or protein function. Analogous technologies are being pursued with peptides⁵¹, carbohydrates⁵² and other small molecules, as well as with cells and tissues⁵³.

The role of materials science in array technologies was demonstrated in striking fashion by Fodor and co-workers⁵⁴, who used photolithographic techniques to prepare an array of 1,024 peptides in just ten sequential operations. Similar methods have been used to create oligonucleotide arrays⁵⁵. Brown and co-workers introduced an alternative approach to DNA arrays by high-speed robotic spotting of complementary DNAs on treated glass surfaces⁵⁶. Both techniques yield 'DNA chips' characterized by high densities of biological information.

Although these technologies are now highly developed and widely used, they are far from optimized. Hybridization of DNA arrays typically requires many hours⁵⁶, precluding their use in circumstances that demand rapid diagnosis or detection of pathogens, and limiting the possibilities for high-throughput screening of large numbers of samples. Array technologies that allow local application of electric fields (to enhance DNA migration)⁵⁷, ultrasonic mixing, or acoustic microstreaming⁵⁸, have all been shown to reduce substantially the time required for read-out of microarray data. Thermal-gradient DNA chips have been developed to allow local (site-to-site) control of the hybridization temperature⁵⁹, which can be important in distinguishing perfect sequence matches from single-site mismatches, and microporous arrays have been used in flow systems to speed up hybridization⁶⁰.

Additional challenges arise in the fabrication of protein arrays. Unlike nucleic acids, which share a common physical chemistry that is largely independent of sequence, proteins are highly diverse in

terms of charge, hydrophobic character, and so on. So whereas robotic printing of DNAs on a common substrate can be done reliably, similar treatment of a complex protein library is unlikely to result in uniform retention of protein structure and function. The potential importance of protein arrays in the direct determination of protein concentrations in cells and tissue fluids, in the identification of protein-protein interactions, and in the study of drug-binding behaviour, has prompted substantial effort directed towards the development of reliable methods for array fabrication⁵⁰.

A successful fabrication method must be applicable to diverse protein libraries, and must achieve robust attachment of each protein to the surface of the array device. The attached proteins must not denature (unfold), and at least some fraction of each protein must be presented in an orientation that is consistent with retention of function; the site(s) of interest cannot be blocked by the array surface or by the linking groups used to achieve surface attachment. The substrate should not be prone to non-specific protein adsorption, because non-specific adsorption can give rise to high background noise and false positive signals.

Brown and co-workers examined the simple approach of printing proteins (either antigens or antibodies) on poly(L-lysine)-coated glass microscope slides⁶¹. They examined 115 antibody/antigen pairs, and found that approximately 50% of the printed antigens yielded signals that were quantitatively consistent with the known concentrations of the corresponding antibodies in the analyte solutions. Antibody arrays performed at a significantly lower level. This study is especially instructive because it provides a careful analysis of the quantitative reliability of the technology, and because it illustrates the challenges involved in fabricating protein arrays. The fact that fewer than 50% of arrayed antibodies yielded reliable signals is particularly striking given the structural uniformity of antibodies. More diverse protein libraries are likely to be substantially more difficult to array.

Alternative approaches to the fabrication of protein microarrays have used aldehydic surfaces to capture protein-bound lysine side chains⁶², polyacrylamide gel pads to entrap proteins and to immobilize antibodies via reaction with oxidized carbohydrate substituents⁶³, and 'capture ligands' that react covalently and specifically with a single 'capture protein' that can be fused to each of the diverse proteins in the library to be arrayed⁶⁴. The most ambitious experi-

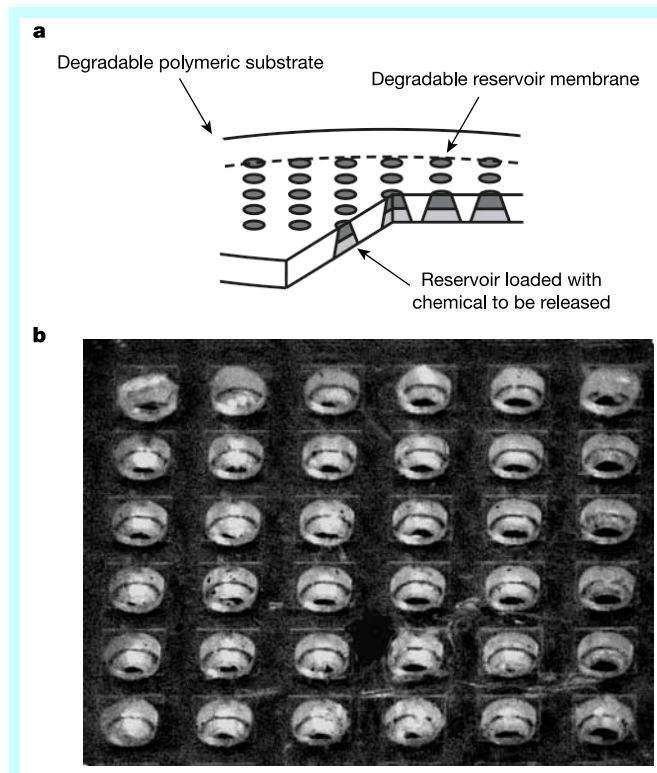


Figure 3 Degradable polymeric microchip. **a**, Cut-away diagram. **b**, Close-up photograph of reservoirs on a degradable microchip composed of polylactic-glycolic acid. Image courtesy of A. Richards-Grayson.

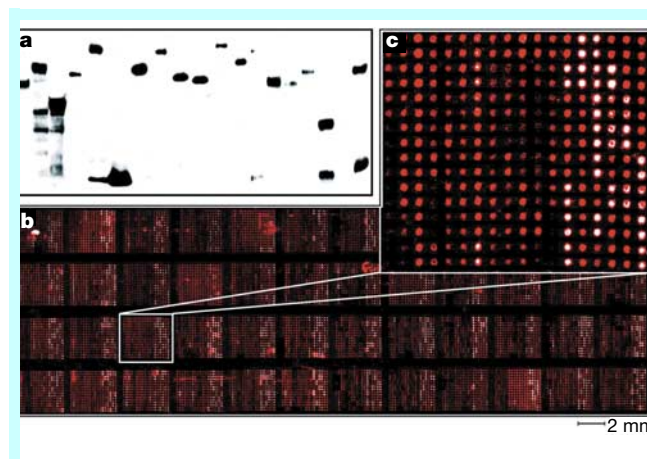


Figure 4 Miroarray of 5,800 yeast proteins, each fused to a hexahistidine sequence that facilitates immobilization on a nickel-coated glass slide. A second fused sequence, that of glutathione *S*-transferase (GST), allows visualization with anti-GST antibodies. Arrays of this kind have been developed by Snyder and co-workers, and used in global analyses of the yeast proteome. **a**, Immunoblot analysis of 19 representative fusion proteins purified in 96-well format; proteins were visualized with anti-GST. **b**, Protein samples spotted on nickel-coated glass slide and probed with anti-GST. **c**, Enlarged image of a portion of the protein array. Image from ref. 65.

ment to date⁶⁵ has used a polyhistidine tag to immobilize more than 5,000 yeast proteins to the surface of a glass slide functionalized with Ni²⁺ (Fig. 4)

Arrayed of membrane proteins presents special challenges. Membrane proteins are especially difficult to isolate and to maintain in active form. A promising approach to membrane protein arrays has emerged from studies of supported lipid bilayers, and from the discovery that supported bilayers can be partitioned into 'corrals' that do not allow lipid intermixing via lateral diffusion⁶⁶. Partitioning can be accomplished by microcontact printing, by deposition of molecular barriers, or by a variety of photochemical methods. The supported bilayer stabilizes membrane proteins with respect to denaturation, and the lateral mobility of the system facilitates protein-protein interactions that are critical to many biological processes.

There is as yet no consensus regarding the preferred method(s) of array fabrication, and further development will be required before protein arrays become widely available for use in research and clinical practice. Materials that reduce non-specific adsorption and protein denaturation, that allow further reduction in feature size, that facilitate address of individual array features, and that enable signal amplification and enhanced sensitivity, will constitute an important part of such development.

Concluding remarks

Biomaterials have already had an enormous impact on health care, and are already widely used in prosthetic and drug delivery devices. Much of this success has been achieved through judicious selection of existing materials, with little real design for biomedical use. Current work is laying the foundations for a much richer application of biomaterials through elucidation of the fundamental design considerations that determine the success or failure of biomaterials systems.

Numerous challenges remain in biomaterials development. These challenges include targeting materials (containing drugs), for example, to specific cells; designing materials that can sense biochemical signals in the body; and developing materials with improved biocompatibility. The ability to address these challenges will be facilitated by advances in biology and materials science. Understanding more about extracellular matrix biology, cell receptors, and immunology will help, for example, in understanding how the body responds to specific materials. Analogously, advances in nanomaterials will create new opportunities to mimic entities in the body (such as cells), and advances in materials characterization will aid in understanding how materials interface with cells and tissues. Finally, we note that many significant recent advances in biomaterials occurred at the interface of clinical medicine and materials science and engineering. Creating opportunities and training programs for cross-disciplinary research for individuals engaged in these areas could significantly accelerate the advancement of biomaterials and create new applications for these materials in medicine. □

doi:10.1038/nature02388.

- Peppas, N. A. & Langer, R. New challenges in biomaterials. *Science* **263**, 1715–1720 (1994).
- Ratner, B. D., Hoffman, A. S., Schoen, J. E. & Lemons, J. E. *Biomaterials Science, an Introduction to Materials in Medicine* 1–8 (Academic, San Diego, 1996).
- Bell, E., Ivarsson, B. & Merrill, C. Production of a tissue-like structure by contraction of collagen lattices by human fibroblasts of different proliferative potential in vitro. *Proc. Natl Acad. Sci. USA* **76**, 1274–1278 (1979).
- Langer, R. Drug delivery and targeting. *Nature* **392** (Suppl.), 5–10 (1998).
- Langer, R. Where a pill won't reach. *Sci. Am.* **288**, 50–57 (2003).
- Morice, M. et al. A randomized comparison of a sirolimus-eluting stent with a standard stent for coronary revascularization. *N. Engl. J. Med.* **346**, 1773–1780 (2002).
- Langer, R. Perspectives: Drug delivery—Drugs on target. *Science* **293**, 58–59 (2001).
- Vacanti, J. P. & Langer, R. Tissue engineering: the design and fabrication of living replacement devices for surgical reconstruction and transplantation. *Lancet* **354**, 32–34 (1999).
- Yurchenco, P. D., Birk, D. E. & Mecham, R. P. (eds) *Extracellular Matrix Assembly and Structure* (Academic, San Diego, 1994).
- van Hest, J. C. M. & Tirrell, D. A. Protein-based materials: Toward a new level of structural control.

Chem. Commun. **19**, 1897–1904 (2001).

- Lee, J., Macosko, C. W. & Urry, D. W. Mechanical properties of cross-linked synthetic elastomeric polypentapeptides. *Macromolecules* **34**, 5968–5974 (2001).
- Nagapudi, K. et al. Photomediated solid-state cross-linking of an elastin-mimetic recombinant protein polymer. *Macromolecules* **35**, 1730–1737 (2002).
- McMillan, R. A. & Conticello, V. P. Synthesis and characterization of elastin-mimetic protein gels derived from a well-defined polypeptide precursor. *Macromolecules* **33**, 4809–4821 (2000).
- Huang, L. et al. Generation of synthetic elastin-mimetic small diameter fibers and fiber networks. *Macromolecules* **33**, 2989–2997 (2000).
- Heilshorn, S. C., DiZio, K. A., Welsh, E. R. & Tirrell, D. A. Endothelial cell adhesion to the fibronectin CS5 domain in artificial extracellular matrix proteins. *Biomaterials* **24**, 4245–4252 (2003).
- Urry, D. W., Parker, T. M., Reid, M. C. & Gowda, D. C. Biocompatibility of the bioelastic material poly(GVGVP) and its γ -irradiation crosslinked matrix. *J. Bioact. Compat. Polym.* **3**, 263–282 (1991).
- Kwon, I., Kirshenbaum, K. & Tirrell, D. A. Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc.* **125**, 7512–7513 (2003).
- Alsberg, E., Anderson, K. W., Albeiruti, A., Rowley, J. A. & Mooney, D. J. Engineering growing tissues. *Proc. Natl Acad. Sci. USA* **99**, 12025–12030 (2002).
- Zhang, S. Emerging biological materials through molecular self-assembly. *Biotechnol. Adv.* **20**, 321–339 (2002).
- Hartgerink, J. D., Beniash, E. & Stupp, S. Peptide-amphiphile nanofibers: A versatile scaffold for the preparation of self-assembling materials. *Proc. Natl Acad. Sci. USA* **99**, 5133–5138 (2002).
- Barrera, D. A., Zylstra, E., Lansbury, P. T. & Langer, R. Synthesis and RGD peptide modification of a new biodegradable copolymer system: Poly(lactic acid-co-lysine). *J. Am. Chem. Soc.* **115**, 11010–11011 (1993).
- Cook, A. D. et al. Characterization and development of RGD-peptide-modified poly(lactic acid-co-lysine) as an interactive, resorbable biomaterial. *J. Biomed. Mater. Res.* **35**, 513–523 (1997).
- Gref, R. et al. Biodegradable long-circulating polymeric nanospheres. *Science* **263**, 1600–1603 (1994).
- Halstenberg, S., Panitch, A., Rizzi, S., Hall, H. & Hubbell, J. A. Biologically engineered protein-graft-poly(ethylene glycol) hydrogels: A cell adhesive and plasmin-degradable biosynthetic material for tissue repair. *Biomacromolecules* **3**, 710–723 (2002).
- Miyata, T., Asami, N. & Urugami, T. A reversibly antigen-responsive hydrogel. *Nature* **399**, 766–768 (1999).
- Kelch, S. & Lendlein, A. Shape memory polymers. *Angew. Chem. Int. Edn Engl.* **41**, 2034–2057 (2002).
- Lendlein, A. & Langer, R. Biodegradable, elastic shape-memory polymers for potential biomedical applications. *Science* **296**, 1673–1676 (2002).
- Pathak, C. P., Swahney, A. S. & Hubbell, J. A. Rapid photopolymerization of immunoprotective gels in contact with cells and tissue. *J. Am. Chem. Soc.* **114**, 8311–8312 (1992).
- Anseth, K., Shastri, V. & Langer, R. Photopolymerizable degradable polyanhydrides with osteocompatibility. *Nature Biotechnol.* **17**, 156–159 (1999).
- Elisseff, J. et al. Transdermal photopolymerization for minimally invasive implantation. *Proc. Natl Acad. Sci. USA* **96**, 3104–3107 (1999).
- Peppas, N. A. Hydrogels and drug delivery. *Curr. Opin. Colloid Interf. Sci.* **2**, 531–537 (1997).
- Lahann, J. et al. A reversible switching of surfaces. *Science* **299**, 371–374 (2003).
- Brocchini, S., James, K., Tangpasuthadol, V. & Kohn, J. Structure-property correlations in a combinatorial library of degradable biomaterials. *J. Biomed. Mater. Res.* **42**, 66–75 (1998).
- Belu, A. M., Brocchini, S., Kohn, J. & Ratner, B. D. Characterization of combinatorially designed polyarylates by time-of-flight secondary ion mass spectrometry. *Rapid Commun. Mass Spectrom.* **14**, 564–571 (2000).
- Lynn, D. M., Anderson, D. G., Putnam, D. & Langer, R. Accelerated discovery of synthetic transfection vectors: Parallel synthesis and screening of degradable polymer library. *J. Am. Chem. Soc.* **123**, 8155–8156 (2001).
- Anderson, D., Lynn, D. & Langer, R. Semi-automated synthesis and screening of a large library of degradable cationic polymers for gene delivery. *Angew. Chem.* **42**, 3153–3158 (2003).
- Luo, D. & Saltzman, W. M. Synthetic DNA delivery systems. *Nature Biotechnol.* **18**, 33–37 (2000).
- Pun, S. H. & Davis, M. E. Development of a nonviral gene delivery vehicle for systemic application. *Bioconjugate Chem.* **13**, 630–639 (2002).
- Affleck, D. G., Yu, L., Bull, D. A., Bailey, S. H. & Kim, S. W. Augmentation of myocardial transfection using TerplexDNA: a novel gene delivery system. *Gene Ther.* **8**, 349–353 (2001).
- McManus, M. T. & Sharp, P. A. Gene silencing in mammals by small interfering RNAs. *Nature Rev. Genet.* **3**, 737–747 (2002).
- Mathiowitz, E. et al. Biologically erodable microspheres as potential oral drug delivery systems. *Nature* **386**, 410–414 (1997).
- Torres-Lugo, M., Garcia, M., Record, R. & Peppas, N. A. pH-sensitive hydrogels as gastrointestinal tract absorption enhancers: Transport mechanisms of salmon calcitonin and other model molecules using the Caco-2 cell model. *Biotechnol. Progr.* **18**, 612–616 (2002).
- Santini, J. T., Cima, M. J. & Langer, R. A controlled-release microchip. *Nature* **397**, 335–338 (1999).
- Grayson, A. et al. Multi-pulse drug delivery from a resorbable polymeric microchip device. *Nature Mater.* **2**, 767–772 (2003).
- Stangel, K. et al. A programmable intraocular CMOS pressure sensor system implant. *IEEE J. Solid-State Circuits* **36**, 1094–1100 (2001).
- Schwartz, M. et al. Single chip CMOS imagers and flexible microelectronic stimulators for a retina implant system. *Sensors Actuators* **83**, 40–46 (2000).
- Kaushik, S. et al. Lack of pain associated with microfabricated microneedles. *Anesth. Analg.* **92**, 502–504 (2001).
- Borenstein, J. T. et al. Microfabrication technology for vascularized tissue engineering. *Biomed. Microdevices* **4**, 671–680 (1999).
- Eisen, M. B. & Brown, P. O. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303**, 179–205 (1999).
- Zhu, H. & Snyder, M. Protein chip technology. *Curr. Opin. Chem. Biol.* **7**, 55–63 (2003).
- Houseman, B. T., Huh, J. H., Kron, S. J. & Mrksich, M. Peptide chips for the quantitative evaluation of protein kinase activity. *Nature Biotechnol.* **20**, 270–274 (2002).
- Houseman, B. T. & Mrksich, M. Carbohydrate arrays for the evaluation of protein binding and enzymatic modification. *Chem. Biol.* **9**, 443–454 (2002).

53. Kononen, J. *et al.* Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nature Med.* **4**, 844–847 (1998).
54. Fodor, S. P. *et al.* Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767–773 (1991).
55. McGall, G. *et al.* Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. *Proc. Natl Acad. Sci. USA* **93**, 13555–13560 (1996).
56. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).
57. Sosnowski, R., Tu, E., Butler, W., O'Connell, J. & Heller, M. Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. *Proc. Natl Acad. Sci. USA* **94**, 1119–1123 (1997).
58. Hui Liu, R., Lenigk, R., Druyor-Sanchez, R. L., Yang, J. & Grodzinski, P. Hybridization enhancement using cavitation microstreaming. *Anal. Chem.* **75**, 1911–1917 (2003).
59. Kajiyama, T. *et al.* Genotyping on a thermal gradient DNA chip. *Genome Res.* **13**, 467–475 (2003).
60. Cheek, B. J., Steel, A. B., Torres, M. P., Yu, Y. & Yang, H. Chemiluminescence detection for hybridization assays on the flow-thru chip, a three-dimensional microchannel biochip. *Anal. Chem.* **73**, 5777–5783 (2001).
61. Haab, B. B., Dunham, M. J. & Brown, P. O. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* **2**, 1–13 (2001).
62. MacBeath, G. & Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763 (2000).
63. Arenkov, P. *et al.* Protein microchips: Use for immunoassay and enzymatic reactions. *Anal. Biochem.* **28**, 123–131 (2000).
64. Hodneland, C. D., Lee, Y.-S., Min, D.-H. & Mrksich, M. Selective immobilization of proteins to self-assembled monolayers presenting active site-directed capture ligands. *Proc. Natl Acad. Sci. USA* **99**, 5048–5052 (2002).
65. Zhu, H. *et al.* Global analysis of protein activities using proteome chips. *Science* **293**, 2101–2105 (2001).
66. Groves, J. T. & Boxer, S. G. Micropattern formation in supported lipid membranes. *Acc. Chem. Res.* **35**, 149–157 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements This work was supported in part by the National Institutes of Health.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.L. (rlanger@mit.edu).

Genome sequence of the Brown Norway rat yields insights into mammalian evolution

Rat Genome Sequencing Project Consortium*

*Lists of participants and affiliations appear at the end of the paper

The laboratory rat (*Rattus norvegicus*) is an indispensable tool in experimental medicine and drug development, having made inestimable contributions to human health. We report here the genome sequence of the Brown Norway (BN) rat strain. The sequence represents a high-quality 'draft' covering over 90% of the genome. The BN rat sequence is the third complete mammalian genome to be deciphered, and three-way comparisons with the human and mouse genomes resolve details of mammalian evolution. This first comprehensive analysis includes genes and proteins and their relation to human disease, repeated sequences, comparative genome-wide studies of mammalian orthologous chromosomal regions and rearrangement breakpoints, reconstruction of ancestral karyotypes and the events leading to existing species, rates of variation, and lineage-specific and lineage-independent evolutionary events such as expansion of gene families, orthology relations and protein evolution.

Darwin believed that "natural selection will always act very slowly, often only at long intervals of time"¹. The consequences of evolution over timescales of approximately 1,000 millions of years (Myr) and 75 Myr were investigated in publications comparing the human with invertebrate and mouse genomes, respectively^{2,3}. Here we describe changes in mammalian genomes that occurred in a shorter time interval, approximately 12–24 Myr (refs 4, 5) since the common ancestor of rat and mouse.

The comparison of these genomes has produced a number of insights:

- The rat genome (2.75 gigabases, Gb) is smaller than the human (2.9 Gb) but appears larger than the mouse (initially 2.5 Gb (ref. 3) but given as 2.6 Gb in NCBI build 32, see <http://www.ncbi.nlm.nih.gov/genome/seq/NCBIContigInfo.html>).
- The rat, mouse and human genomes encode similar numbers of genes. The majority have persisted without deletion or duplication since the last common ancestor. Intronic structures are well conserved.
- Some genes found in rat, but not mouse, arose through expansion of gene families. These include genes producing pheromones, or involved in immunity, chemosensation, detoxification or proteolysis.
- Almost all human genes known to be associated with disease have orthologues in the rat genome but their rates of synonymous substitution are significantly different from the remaining genes.
- About 3% of the rat genome is in large segmental duplications, a fraction intermediate between mouse (1–2%) and human (5–6%). These occur predominantly in pericentromeric regions. Recent expansions of major gene families are due to these genomic duplications.
- The eutherian core of the rat genome—that is, bases that align orthologously to mouse and human—comprises a billion nucleotides (~40% of the euchromatic rat genome) and contains the vast majority of exons and known regulatory elements (1–2% of the genome). A portion of this core constituting 5–6% of the genome appears to be under selective constraint in rodents and primates, while the remainder appears to be evolving neutrally.
- Approximately 30% of the rat genome aligns only with mouse, a considerable portion of which is rodent-specific repeats. Of the non-aligning portion, at least half is rat-specific repeats.
- More genomic changes occurred in the rodent lineages than the

primate: (1) These rodent genomic changes include approximately 250 large rearrangements between a hypothetical murid ancestor and human, approximately 50 from the murid ancestor to rat, and about the same from the murid ancestor to mouse. (2) A threefold-higher rate of base substitution in neutral DNA is found along the rodent lineage when compared with the human lineage, with the rate on the rat branch 5–10% higher than along the mouse branch. (3) Microdeletions occur at an approximately twofold-higher rate than microinsertions in both rat and mouse branches.

- A strong correlation exists between local rates of microinsertions and microdeletions, transposable element insertion, and nucleotide substitutions since divergence of rat and mouse, even though these events occurred independently in the two lineages.

Background

History of the rat

The rat, hated and loved at once, is both scourge and servant to mankind. The "Devil's Lapdog" is the first sign in the Chinese zodiac and traditionally carries the Hindu god Ganesh⁶. Rats are a reservoir of pathogens, known to carry over 70 diseases. They are involved in the transmission of infectious diseases to man, including cholera, bubonic plague, typhus, leptospirosis, cowpox and hantavirus infections. The rat remains a major pest, contributing to famine with other rodents by eating around one-fifth of the world's food harvest.

Paradoxically, the rat's contribution to human health cannot be overestimated, from testing new drugs, to understanding essential nutrients, to increasing knowledge of the pathobiology of human disease. In many parts of the world the rat remains a source of meat.

The laboratory rat (*R. norvegicus*) originated in central Asia and its success at spreading throughout the world can be directly attributed to its relationship with humans⁷. J. Berkenhout, in his 1769 treatise *Outline of the Natural History of Great Britain*, mistakenly took it to be from Norway and used *R. norvegicus* Berkenhout in the first formal Linnaean description of the species. Whereas the black rat (*Rattus rattus*) was part of the European landscape from at least the third century AD and is the species associated with the spread of bubonic plague, *R. norvegicus* probably originated in northern China and migrated to Europe somewhere

around the eighteenth century⁸. They may have entered Europe after an earthquake in 1727 by swimming the Volga river.

The rat in research

R. norvegicus was the first mammalian species to be domesticated for scientific research, with work dating to before 1828 (ref. 9). The first recorded breeding colony for rats was established in 1856 (ref. 9). Rat genetics had a surprisingly early start. The first studies by Crampe from 1877 to 1885 focused on the inheritance of coat colour¹⁰. Following the rediscovery of Mendel's laws at the turn of the century, Bateson used these concepts in 1903 to demonstrate that rat coat colour is a mendelian trait¹⁰. The first inbred rat strain, PA, was established by King in 1909, the same year that systematic inbreeding began for the mouse¹⁰. Despite this, the mouse became the dominant model for mammalian geneticists, while the rat became the model of choice for physiologists, nutritionists and other biomedical researchers. Nevertheless, there are over 234 inbred strains of *R. norvegicus* developed by selective breeding, which 'fixes' natural disease alleles in particular strains or colonies¹¹.

Over the past century, the role of the rat in medicine has transformed from carrier of contagious diseases to indispensable tool in experimental medicine and drug development. Current examples of use of the rat in human medical research include surgery¹², transplantation^{13–15}, cancer^{16,17}, diabetes^{18,19}, psychiatric disorders²⁰ including behavioural intervention²¹ and addiction²², neural regeneration^{23,24}, wound^{25,26} and bone healing²⁷, space motion sickness²⁸, and cardiovascular disease^{29–31}. In drug development, the rat is routinely employed both to demonstrate therapeutic efficacy^{15,32,33} and to assess toxicity of novel therapeutic compounds before human clinical trials^{34–37}.

The Rat Genome Project

Over the past decade, investigators and funding agencies have participated in rat genomics to develop valuable resources. Before the launch of the Rat Genome Sequencing Project (RGSP), there was much debate about the overall value of the rat genome sequence and its contribution to the utility of the rat as a model organism. The debate was fuelled by the naive belief that the rat and mouse were so similar morphologically and evolutionarily that the rat sequence would be redundant. Nevertheless, an effort spearheaded by two NIH agencies (NHGRI and NHLBI) culminated in the formation of the RGSP Consortium (RGSPC).

The RGSP was to generate a draft sequence of the rat genome, and, unlike the comparable human and mouse projects, errors would not ultimately be corrected in a finished sequence³⁸. Consequently, the draft quality was critical. Although it was expected to have gaps and areas of inaccuracy, the overall sequence quality had to be high enough to support detailed analyses.

The BN rat was selected as a sequencing target by the research community. An inbred animal (BN/SsNHsd) was obtained by the Medical College of Wisconsin (MCW) from Harlan Sprague Dawley. Microsatellite studies indicated heterozygosity, so over 13 generations of additional inbreeding were performed at the MCW, resulting in BN/SsNHsd/Mcwi animals. Most of the sequence data were from two females, with a small amount of whole genome shotgun (WGS) and flow-sorted Y chromosome sequencing from a male. The Y chromosome is not included in the current assembly.

A network of centres generated data and resources, led by the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC) and including Celera Genomics, the Genome Therapeutics Corporation, the British Columbia Cancer Agency Genome Sciences Centre, The Institute for Genomic Research, the University of Utah, the Medical College of Wisconsin, The Children's Hospital of Oakland Research Institute, and the Max Delbrück Center for Molecular Medicine, Berlin. After assembly of the genome at the

BCM-HGSC, analysis was performed by an international team, representing over 20 groups in six countries and relying largely on gene and protein predictions produced by Ensembl.

Determination of the genome sequence

Atlas and the 'combined' sequencing strategy

Despite progress in assembling draft sequences^{2,3,39–44} the question of which method produces the highest-quality products is unresolved. A significant issue is the choice between logistically simpler WGS approaches versus more complex strategies employing bacterial artificial chromosome (BAC) clones^{45–48}. In the Public Human Genome Project² a BAC by BAC hierarchical approach was used and provided advantages in assembling difficult parts of the genome. The draft mouse sequence was a pure WGS approach using the ARACHNE assembler^{3,49,50} but underrepresented duplicated regions owing to 'collapses' in the assembly^{3,51–53}. This limitation of the mouse draft sequence was tolerable owing to the planned full use of BAC clones in constructing the final finished sequence.

The RGSPC opted to develop a 'combined' approach using both WGS and BAC sequencing (Fig. 1). In the combined approach, WGS data are progressively melded with light sequence coverage of individual BACs (BAC skims) to yield intermediate products called 'enriched BACs' (eBACs). eBACs covering the whole genome are then joined into longer structures (bactigs). Bactigs are joined to form larger structures: superbactigs, then ultrabactigs. During this process other data are introduced, including BAC end sequences, DNA fingerprints and other long-range information (genetic markers, syntenic information), but the process is constrained by eBAC structures.

To execute the combined strategy we developed the *Atlas* software package⁵⁴ (Fig. 1). The *Atlas* suite includes a 'BAC-Fisher' component that performs the functions needed to generate eBACs. WGS genome coverage was generated ahead of complete BAC coverage, so a BAC-Fisher web server was established at the BCM-HGSC to enable users to access the combined BAC and WGS reads as each BAC was processed (see Methods for data access). Each eBAC is assembled with high stringency to represent the local sequence accurately, and so provide a valuable intermediate product that assists all users of the genome data. Additional *Atlas* modules joined eBACs and linked bactigs to give the complete assembly (Fig. 1). Overall, the combined approach takes advantage of the strengths of both previous methods, with few of the disadvantages.

Sequence and genome data

Over 44 million DNA sequence reads were generated (Table 1; Methods). Following removal of low-quality reads and vector contaminants, 36 million reads were used for *Atlas* assembly, which retained 34 million reads. This was 7× sequence coverage with 60% provided by WGS and 40% from BACs. Slightly different estimates came from considering the entire 'trimmed' length of the sequence data (7.3×), or only the portion of Phred20 quality or higher (6.9×).

The sequence data were end-reads from clones either derived directly from the genome (insert sizes of <10 kb, 10 kb, 50 kb and >150 kb) or from small insert plasmids subcloned from BACs. Overall, these provided 42-fold clone coverage, with 32-fold coverage having both paired ends represented. Approximately equal contributions of clone coverage were from the different categories.

Over 21,000 BACs were used for BAC skims (1.6× coverage) with an average sequence depth of 1.8×, giving an overall 2.8× genomic sequence coverage from BACs. This was slightly more than the most efficient procedure would require (~1.2× each), because the genome size was not known at the project start.

Simultaneous with sequencing, 199,782 clones from the CHORI-230 BAC library⁵⁵ were fingerprinted by restriction enzyme

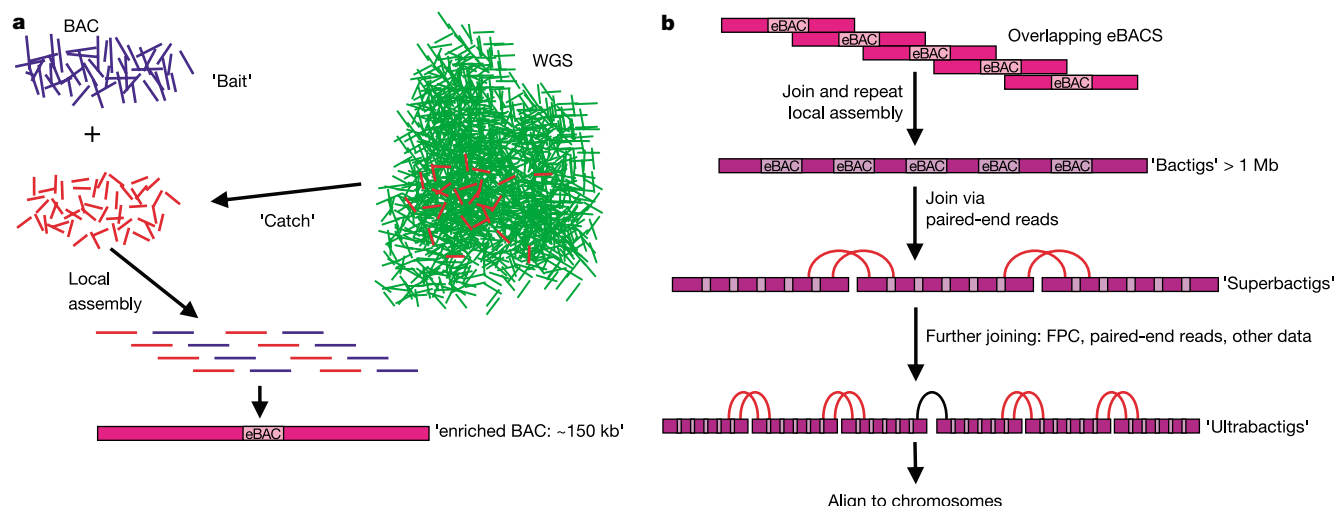


Figure 1 The new 'combined' sequence strategy and *Atlas* software. **a**, Formation of 'eBACs'. The RGSP strategy combined the advantages of both BAC and WGS sequence data²⁴. Modest sequence coverage (~1.8-fold) from a BAC is used as 'bait' to 'catch' WGS reads from the same region of the genome. These reads, and their mate pairs, are assembled using Phrap to form an eBAC. This stringent local assembly retains 95% of the 'catch'. **b**, Creation of higher-order structures. Multiple eBACs are assembled into bactigs

based on sequence overlaps. The bactigs are joined into superbactigs by large clone mate-pair information (at least two links), extended into ultrabactigs using additional information (single links, FPC contigs, synteny, markers), and ultimately aligned to genome mapping data (radiation hybrid and physical maps) to form the complete assembly.

digestion, representing 12-fold genomic coverage⁵⁶ (Methods). These were assembled into a 'fingerprint contig (FPC)' map (a contig is a set of overlapping segments of DNA) containing 11,274 FPCs. BAC selection for sequence skimming was based on overlaps between BACs using FPC mapping⁵⁶ (M.K. and C.F., unpublished work), ongoing BAC end sequencing (S.Z., unpublished work), and BAC sequence skimming⁵⁷. This strategy led to the sequence of a tiling path of BAC clones, covering the whole genome. In addition to the FPC map, a yeast artificial chromosome (YAC)-based physical map was constructed. 5,803 BAC and P1-derived artificial chromosome (PAC) clones from RPCI-32 and RPCI-31 libraries⁵⁵, respectively, were anchored to 51,323 YAC clones originating from two tenfold-coverage YAC libraries^{58,225} assembled into 605 contigs⁵⁶. This map was subsequently integrated with the FPC map and the sequence assembly, reducing the total number of map contigs to 376 (minimum length of contig containing the 'typical' nucleotide, $N_{50} = 172$ clones, 4.4 Mb; 358 anchored to the sequence assembly; Supplementary Information).

The combined strategy enabled development of resources such as the FPC map, BAC end sequences, and BAC skim sequences in parallel, rather than sequentially. In addition to allowing ongoing

quality checking, this permitted the data-gathering phase of the project to be completed in less than two years.

Atlas assembly

Statistics for the Rnor3.1 assembly are in Table 2. Contigs within eBACs were ordered and oriented using read-pair information. Read-pair information was also used to add WGS reads to eBACs, even when sequence overlaps could not be reliably detected owing to repeated sequences. BAC skim reads with repeats were included in the assembly of eBACs because they clearly originated within BAC insert sequences. Over 19,000 eBACs were eventually generated.

More than 98% of eBACs were successfully merged to form bactigs (Fig. 1). Bactigs were subsequently reassembled to process all reads from overlapping BACs simultaneously, and then ordered and oriented with respect to each other using FPC map and BAC end sequence read-pair information. These superbactig and ultrabactig structures (see below) were aligned with chromosomes using external information, such as positions of genetic markers. Ultrabactigs represented the largest sequence units used to build chromosomes.

The current release of the rat genome assembly, version Rnor3.1,

Table 1 Clones and reads used in the RGSP

Insert size* (kb)	Source or vector	Reads (millions)				Bases (billions)		Sequence coverage†		Clone coverage‡
		All§	Used	Paired	Assembled	Trimmed	≥Phred20	Trimmed	≥Phred20	
2–4	Plasmid	9.6	8.6	7.4	7.9	4.8	4.5	1.8	1.6	3.70
4.5–7.5	Plasmid	4.5	4.3	3.6	3.6	2.4	2.3	0.87	0.82	2.96
10	Plasmid	8.4	7.2	6.4	6.4	4.1	3.8	1.5	1.4	11.63
50	Plasmid	1.7	1.3	1.0	1.1	0.69	0.65	0.25	0.24	9.47
150–250	BAC	0.32	0.31	0.26	0.26	0.18	0.16	0.07	0.06	9.26
Total WGS		24.5	21.7	18.7	19.2	12.1	11.3	4.4	4.1	37.0
2–5	BAC skims	19.6	14.6	13.2	14.5	8.0	7.7	2.9	2.8	4.8
Total		44.1	36.3	31.9	33.7	20.2	19.0	7.3	6.9	41.8

* Grouped in ranges of sizes for individual libraries tracked to specific multiples of 0.5 kb.

† Total bases in used reads divided by sampled genome size including all cloned and sequenced euchromatic or heterochromatic regions.

‡ Estimated as sum of insert sizes divided by sampled genome size.

§ WGS reads available on the NCBI Trace Archive as of 21 March 2003; BAC skim reads attempted at BCM-HGSC as of 12 May 2003; BAC end reads obtained directly from TIGR.

|| Refers to coverage from 2–5 kb subclones from BACs. The BACs that were skimmed amounted to 1.58 × clone coverage.

Table 2 Statistics of the RGSP draft sequence assembly

Features*	Number	N50 length (kb)	Bases (Gb)	Bases plus gaps† (Gb)	Percentage of genome‡			
					Sampled (2.78 Gb)		Assembled (2.75 Gb)	
Anchored contigs	127,810	38	2.476	2.481	Bases	Bases + gaps	Bases	Bases + gaps
Anchored superbactig scaffolds	783	5,402	2.476	2.509	89.1	89.2	90.0	90.2
Anchored ultrabactigs	291	18,985	2.476	2.687	89.1	96.6	90.0	97.7
Unanchored superbactigs, main scaffolds	134	1,210	0.056	0.062	2.0	2.2	2.0	2.3
Unanchored ultrabactigs	128	1,529	0.056	0.069	2.0	2.5	2.0	2.5
All superbactigs, main scaffolds	917	5,301	2.533	2.571	91.1	92.5	92.1	93.5
Minor scaffolds	4,345	8	0.033	0.038	1.2	1.4	1.2	1.4

* Anchored sequences are those that can be placed on chromosomes because they contain known markers. The main scaffold for each superbactig is the largest set of contigs (in terms of total contig sequence) that can be ordered and oriented using mate-pair links and ordering of BACs. Scaffolds that cannot be ordered and oriented with respect to the main scaffold are termed minor scaffolds.

† Ambiguous bases (N) are counted in the gap sizes, and excluded in the base counts.

‡ Computed as bases plus gaps divided by estimated genome size. Sampled genome size is based on oligonucleotide frequency statistics of unassembled WGS reads. Assembled genome size is based on cumulative contig sequence following assembly.

was generated using the data in Table 1. Earlier releases (Rnor2.0/2.1, Methods) were used for a substantial part of the annotation and analysis of genes and proteins, whereas the current release provided the genome description. Rnor3.1 has 128,000 contigs, with N_{50} length 38 kb—larger than the expected genomic extent of a mammalian gene. These sequence contigs were linked into 783 superbactigs that were anchored to the radiation hybrid map⁵⁹. These larger units had N_{50} length 5.4 Mb. Another 134 smaller superbactigs (N_{50} length 1.2 Mb) could not be anchored, presumably because they fell into gaps between markers or because they were in repeated regions that could not be unambiguously placed. From placement on the radiation hybrid map, adjacent superbactigs were further linked to maximize continuity of sequence if appropriate read-pair mates existed or FPC suggested links. This reduced linked superbactigs to 419 pieces with 71 singletons. 291 ultrabactigs with N_{50} length of nearly 19 Mb were placed on chromosomes. Orthology information with mouse and human sequences was also used to resolve conflicts and suggest placement of sequence units. Most of the 128 unplaced units were either singletons or small superbactigs that consisted of few clones. Thus, nearly the entire genome was represented in less than 300 large sequence units.

Quality assessment

Thirteen megabases of high-quality finished rat sequence from BACs were available for comparison with Rnor3.1 (Methods). This analysis showed that the majority of draft bases from within contigs were high quality (1.32 mismatches per 10 kb). This is essentially the accepted accuracy standard for finished sequence (1.0 errors per 10 kb)⁶⁰, so the overwhelming majority of contig bases are highly accurate. The highest frequency of mismatches occurred at the ends of contigs. We calculate the average size of these lower-accuracy regions to be 750 base pairs (bp) and they amount to less than 0.9% of the genome. These regions arise from misassembly of terminal reads due to repeated sequences.

Few mismatches were found within contigs. Six were found within contigs when compared with the 13 Mb of finished sequence, or one case per 2.2 Mb. All were insertions or deletions and may represent polymorphisms. Thus, at the fine structure level, the bulk of sequences that make up contigs is nearly the quality of finished sequence.

We judged accuracy of assembly at the chromosomal level by alignment with linkage maps⁶¹ and radiation hybrid map⁵⁹ (Fig. 2). Thirteen markers out of 3,824 from the SHRSP × BN genetic map were placed on different chromosomes in the assembly and in the genetic map. Similarly, of the 20,490 sequence tagged sites placed on both the assembly and radiation hybrid (v3.4) map, 96.9% had consistent chromosome placement⁵⁹. Initial alignments identified regions of misassembly, and these were corrected, so that in Rnor3.1 the maps are congruent except for possible mismapped markers. The distribution of assembled sequence among the chromo-

somes and chromosome sizes in Rnor3.1 are in Supplementary Table SI-2.

Landscape and evolution of the rat genome

Genome size

Genomic assemblies are usually smaller than the actual genome size owing to under-representation of sequences affected by cloning bias, and sequencing and assembly difficulties. Simply equating the assembled genome size with the euchromatic, cloneable portion does not take into account heterochromatin that may be included⁶². We therefore estimated both an assembled genome size, scaled by the inverse of the fraction of features (genetic markers, expressed sequence tags (ESTs), and so on) found in the Rnor3.1 assembly, and a cloneable (or sampled) genome size, which was the part of the genome present in the WGS reads before assembly, as measured by analysing the distribution of short oligomers⁶³. The former may be an underestimate because non-repetitive, easily assembled regions can be enriched for known features. The latter should be an overestimate because there are likely to be regions (such as repeats) that can be cloned and sequenced, but not assembled.

For the rat genome, the assembled and cloneable genome sizes are very close. Considering the fraction of the marker set successfully mapped to Rnor3.1 (92%), or the fraction of sequence finished outside the BCM-HGSC (to reduce bias) present in Rnor3.1 (91%), together with the assembled bases in main scaffolds (2.533 Gb, Table 2), we suggest a genome size of 2.75 Gb. Alternatively, analysis of the WGS oligomers of length 24 to 32 predicted a genome size of between 2.76 and 2.81 billion bases. We have used the more conservative value of 2.75 Gb for the rat genome size, but this is

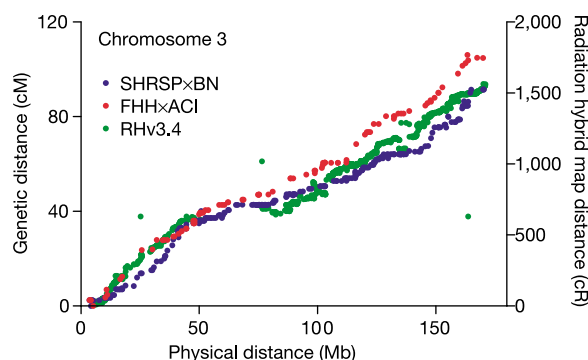


Figure 2 Map correspondence. Correspondence between positions of markers on two genetic maps of the rat (SHRSP × BN intercross and FHH × ACI intercross⁶¹), on the rat radiation hybrid map⁵⁹, and their position on the rat genome assembly (Rnor3.1).

still considerably higher (150 Mb) than the 2.6 Gb currently reported for the mouse draft genome sequence. A fraction of the size differences in these rodent genomes results from the different repeat content (see below); however, it is also recognized that segmental duplications may be under-represented in the mouse WGS draft sequence for technical reasons^{3,51}.

Telomeres, centromeres and mitochondrial sequence

The rat has both metacentric and telocentric chromosomes, in contrast to the wholly telocentric mouse chromosomes. As expected from previous draft sequences, the rat draft does not contain complete telomeres or centromeres. Their physical location relative to the rat draft sequence can however be approximated; the centromeres of the telocentric rat chromosomes (2, 4–10 and X) must be positioned before nucleotide 1 of these assemblies, and those for the remaining chromosomes are estimated as indicated in Fig. 3. Several of these putative centromere positions coincide with both segmental duplication blocks (see below) and classical satellite

clusters, consistent with enrichment of both of these sequence features in rat pericentromeric DNA. Human subtelomere regions are characterized by both an abundance of segmentally duplicated DNA and an enrichment of internal (TTAGGG)_n-like sequence islands⁶⁴. Approximately one-third of the euchromatic rat subtelomeric regions are similarly enriched, suggesting that Rnor3.1 might extend very close to the chromosome ends.

Fragments of the rat mitochondrial genome were also propagated within the WGS libraries and subsequently sequenced, allowing the assembly of the complete 16,313 bp mitochondrial genome (Supplementary Information). Comparison with existing mitochondrial sequences in the public databases revealed variable positions totaling 95 bp (0.6%) between this strain and the wild brown rat. Considerably more variation (2.2%) was found when compared with the Wistar strain: 357 bp differences over the whole genome, including 78 positions that are conserved in the other mammalian sequences. Such variation has also been reported in mouse mitochondrial sequences and attributed to errors in previously

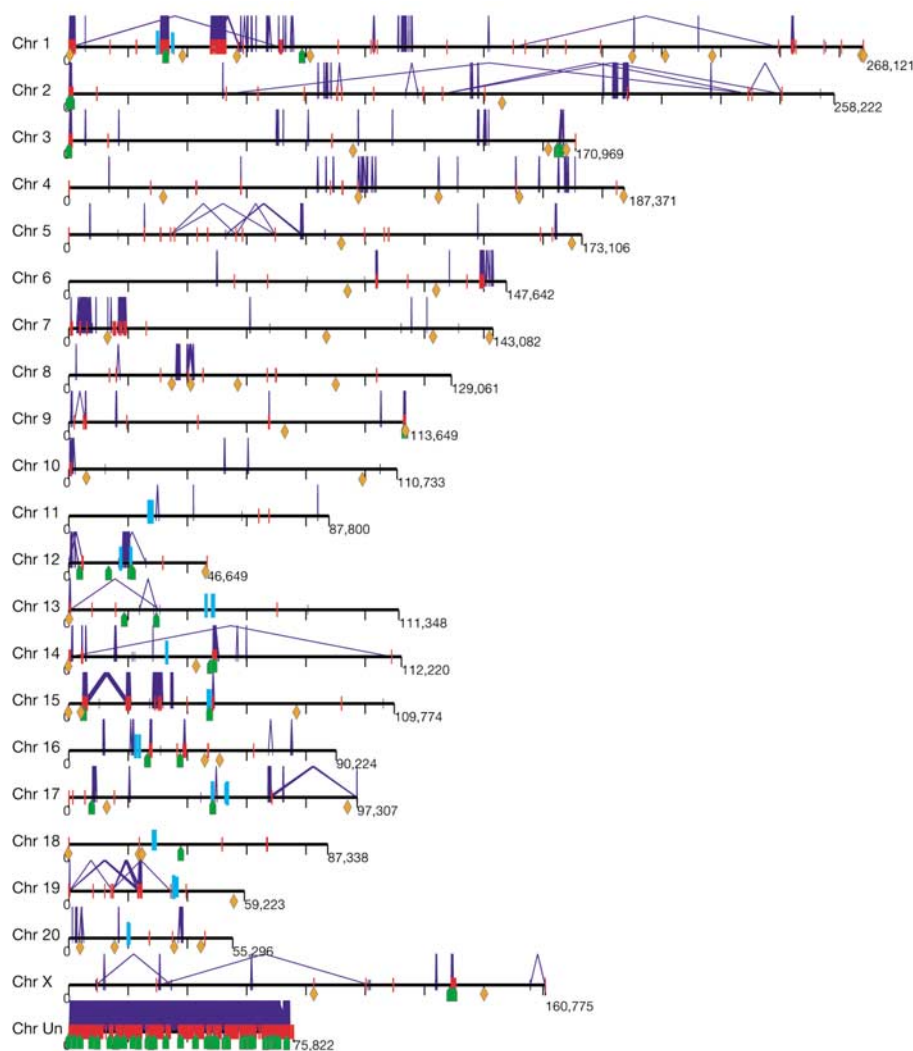


Figure 3 Distribution of segmental duplications in the rat genome. Interchromosomal duplications (red) and intrachromosomal duplications (blue) are depicted for all duplications with $\geq 90\%$ sequence identity and ≥ 20 kb length. The intrachromosomal duplications are drawn with connecting blue line segments; those with no apparent connectors are local duplications (spaced below the figure resolution limit). p arms are on the left and the q arms on the right. Chromosomes 2, 4–10, and X are telocentric; the assemblies begin with pericentric sequences of the q arms, and no centromeres are indicated. For the remaining chromosomes, the approximate centromere positions were

estimated from the most proximal STS/gene marker to the p and q arm as determined by fluorescent *in situ* hybridization (FISH) (cyan vertical lines; no chromosome 3 data). The 'Chr Un' sequence consists of contigs not incorporated into any chromosomes. Green arrows indicate 1 Mb intervals with more than tenfold enrichment of classic rat satellite repeats within the assembly. Orange diamonds indicate 1 Mb intervals with more than tenfold enrichment of internal (TTAGGG)_n-like sequences. For more detail see <http://ratparalogy.cwru.edu>.

sequenced genomes⁶⁵. The current sequence is very accurate, and we therefore favour the BN sequence as a reference for the rat mitochondrial genome.

Orthologous chromosomal segments and large-scale rearrangements

Multi-megabase segments of the chromosomes of the primate–rodent ancestor have been passed on to human and murid rodent descendants with minimal rearrangements of gene order^{66–68}. These intact regions, which are bounded by the breaks that occurred during ancient large-scale chromosomal rearrangements, are referred to as orthologous chromosomal segments. The same phenomenon has occurred in the descent of the rat and mouse from the genome of their common murid ancestor, and we were able to use the human genome, and in some cases other outgroup data, to tentatively reconstruct the sequence of many of these rearrangements in these lineages. To visualize the extent of orthologous chromosomal segments, each genome was ‘painted’ with the orthologous segments of the other two species (Fig. 4) using the Virtual Genome Painting method (M.L.G.-G. *et al.*, unpublished work; <http://www.genboree.org>). Inspection shows the interleaving of events that both preceded and occurred subsequently to the rat–mouse divergence.

Comparing the three species at 1 Mb resolution, BLASTZ⁶⁹, PatternHunter/Grimm-Synteny^{70,71}, Pash⁷², and associated merging algorithms^{66,72,73} produce virtually indistinguishable sets of orthologous chromosomal segments. PatternHunter and the GRIMM-Synteny algorithm⁷³ detect 278 orthologous segments between human and rat, and 280 between human and mouse. The mouse–rat comparison reveals a smaller number of segments (105) of larger average size. The larger number of breaks in orthologous segments between the human to the rodent pair is expected, because of the

latter’s closer evolutionary relationship.

Understanding the number and timing of rearrangement events that have occurred in each of the three individual lineages (see tree in Fig. 5a) since the common primate–rodent ancestor required a more detailed analysis. We initially focused on the X chromosome, because rearrangements between the X and the autosomes are rare⁷⁴ and its history is somewhat easier to trace completely. The X chromosome consists of 16 human–mouse–rat orthologous segments of at least 300 kb in size⁷³ (Fig. 6a). In the most parsimonious scenario (found with MGR and GRIMM⁷⁵), these were created by 15 inversions in the descent from the primate–rodent ancestor (Fig. 6b). Outgroup data from cat, cow⁷⁶ and dog⁷⁷ resolved the timing of these rearrangements more precisely. Most of these events occurred in the rodent lineage: five (or four) before the divergence of rat and mouse, five in the rat lineage, and five in the mouse lineage. At most one rearrangement occurred in the human lineage since divergence from the common ancestor with rodents. The timing of this one event was ambiguous, owing to the limited resolution of the outgroup data. Even given this uncertainty, it is clear that the large-scale architecture of the X chromosome in humans is largely unchanged since the primate–rodent ancestor⁷³, whereas there has been considerable activity in the rodents. The assignment of the accelerated activity to the rodent branch, following the primate–rodent divergence, is consistent with previous studies at significantly lower resolution (these showed complete conservation of marker order between the X chromosomes of human and cat⁷⁸, human and dog⁷⁷, and human and lemur⁷⁹, as well as similar karyotypes of the X chromosomes in human, chimpanzees, gorillas and orangutans⁸⁰).

Large-scale reconstruction of the entire ancestral murid genome suggests that it retained many previously postulated chromosome associations of the placental ancestor^{81,82}. The most parsimonious scenario we found requires a total of 353 rearrangements: 247 between the murid ancestor and human, 50 from the murid ancestor to mouse and 56 from the murid ancestor to rat. A recent study⁸² implies that most of the 247 rearrangements between the murid ancestor and human occurred on the evolutionary subpath from the squirrel–mouse–rat ancestor to the murid ancestor. Our

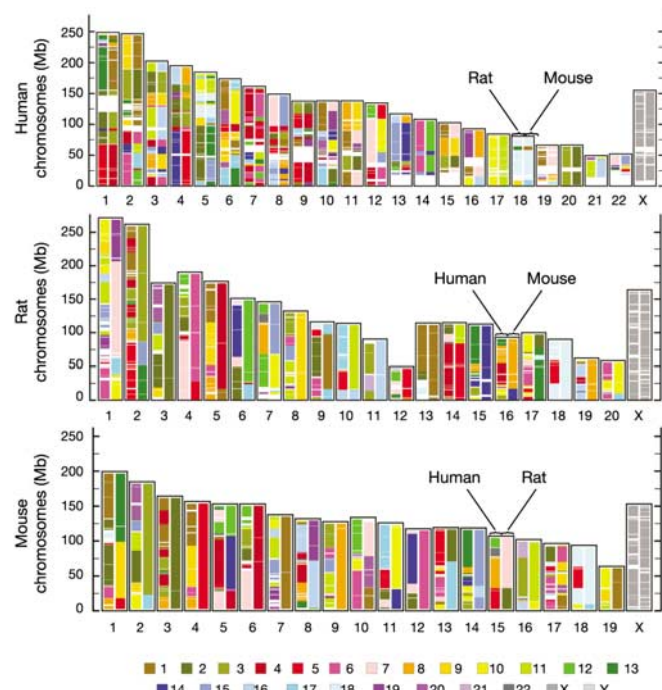
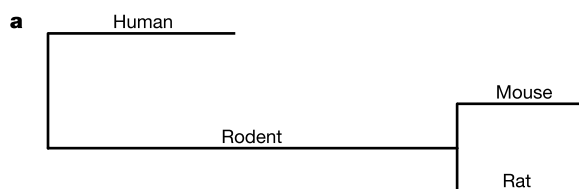


Figure 4 Map of conserved synteny between the human, mouse and rat genomes. For each species, each chromosome (x axis) is a two-column boxed pane (p arm at the bottom) coloured according to conserved synteny to chromosomes of the other two species. The same chromosome colour code is used for all species (indicated below). For example, the first 30 Mb of mouse chromosome 15 is shown to be similar to part of human chromosome 5 (by the red in left column) and part of rat chromosome 2 (by the olive in right column). An interactive version is accessible (<http://www.genboree.org>).



	Substitutions, insertions and deletions			
	Human	Rodent	Mouse	Rat
Substitutions per site	0.11 ±0.0012	0.24 ±0.0012	0.073 ±0.0014	0.077 ±0.006
Substitutions in neutral sites only	0.13 ±0.011	0.28 ±0.033	0.083 ±0.013	0.091 ±0.011
Insertion events per kb	2.7 ±0.94	4.74 ±1.0	1.54 ±0.84	1.43 ±0.73
Deletion events per kb	5.3 ±0.55	12 ±1.2	3.8 ±0.21	4.5 ±0.13
Inserted bases per kb	6.4 ±2.9	9.4 ±1.6	3.6 ±1.5	3.2 ±1.3
Deleted bases per kb	18 ±2.0	40 ±4.9	11 ±0.55	13 ±0.05

Figure 5 Substitutions and microindels (1–10 bp) in the evolution of the human, mouse and rat genomes. **a**, The lengths of the labelled branches in the tree are proportional to the number of substitutions per site inferred using the REV model²²² from all sites with aligned bases in all three genomes. **b**, The table shows the midpoint and variation in these branch-length estimates when estimated from different sequence alignment programs and different neutral sites, including sites from ancestral repeats³, fourfold degenerate sites in codons, and rodent-specific sites (‘in neutral sites only’ row; Supplementary Information). Other rows give midpoints and variation for micro-indels on each branch of the tree in **a**.

analyses confirm that the rate of rearrangements in murid rodents is much higher than in the human lineage⁷³.

Segmental duplications

Segmental duplications are defined here as regions of the genome that are repeated over at least 5 kb of length and >90% identity. The rat has approximately 2.9% of its bases in these duplicated regions (Fig. 3), whereas the human genome has 5–6%⁸³. In contrast to the greater rate of large-scale rearrangement, the mouse genome shows substantially fewer of these events³, with only 1.0–2.0%⁵¹ of its sequenced bases in duplicated regions. These duplicated structures are particularly challenging to assemble, and we attribute at least some of the mouse–rat differences to the BAC-based approach we used for Rnor3.1, compared with the WGS mouse approach. The vast majority of these sequences (73 of 82 Mb) were regions with <99.5% identity and thus were not simply overlapping sequences that had not been joined by the assembly program Phrap. The ‘unplaced’ chromosome in Rnor3.1 showed a marked enrichment for blocks of segmental duplication (nearly 44% of the total), which indicates problems with anchoring these elements to the genome.

Intrachromosomal duplications are represented at a three-to-one excess when compared with interchromosomal duplications, and are significantly enriched near the telomeres and in centromeric regions (Fig. 3). The pericentromeric accumulation of segmental duplications in the rat is reminiscent of that observed in human and mouse^{83–86}, and seems to be a general property of mammalian chromosome architecture.

We observed considerable clustering of duplications⁸⁷, including 41 discrete genomic regions larger than 1 Mb in size in which duplications appear to be organized into groups with <100 kb

between duplicated segments. For many of these clusters, the underlying sequence alignments showed a wide range in the degree of sequence identity, suggesting that these areas have been subject to duplication events more or less continuously over millions of years. In contrast, an analysis of the evolutionary distance between all duplicated regions showed an unusual bimodal distribution, particularly for intrachromosomal segmental duplications. Two peaks were observed at 0.045 substitutions per site and 0.075 substitutions per site. Given that the rat genome has accumulated 8–10% substitutions (see below) since the speciation from mouse 12–24 Myr ago, this bimodal distribution may correspond to bursts of segmental duplication that occurred approximately 5 and 8 Myr ago, respectively.

The segmental duplications in the rat genome were of considerable interest because they represent an important mechanism for the generation of new genes. We found that 63 NCBI reference sequence⁸⁸ (RefSeq; see <http://www.ncbi.nih.gov/RefSeq/>) genes were located completely or partially within rat duplicated regions, out of a genome total of 4,532 rat RefSeq genes. As discussed below, many of these genes are present in multiple copies and belong to gene families that have been recently duplicated and contribute to distinctive elements of rat biology.

Gains and losses of DNA

In addition to large rearrangements and segmental duplications, genome architecture is strongly influenced by insertion and deletion events that add and remove DNA over evolutionary time. To characterize the origins and losses of sequence elements in the human, mouse and rat genomes, we categorized all the nucleotides in each of the three genomes, using our alignment data and

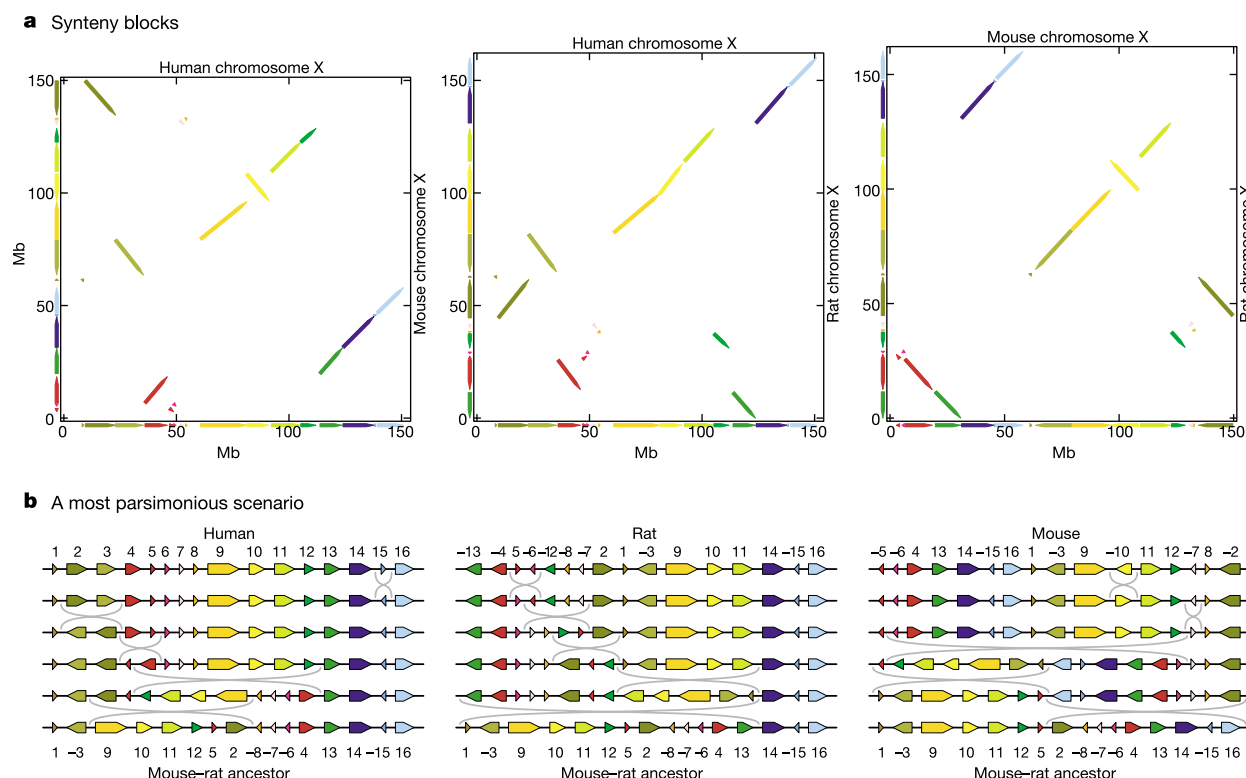


Figure 6 X chromosome in each pair of species. **a**, GRIMM-Synteny⁷¹ computes 16 three-way orthologous segments (≥ 300 kb) on the X chromosome of human, mouse and rat, shown for each pair of species, using consistent colours. **b**, The arrangement (order and orientation) of the 16 blocks implies that at least 15 rearrangement events occurred during X chromosome evolution of these species. The program MGR (<http://www.cs.ucsd.edu/groups/bioinformatics/MGR/>) determined that evolutionary scenarios

with 15 events are achievable and all have the same median ancestor (located at the last common mouse–rat ancestor). Shown is a possible (not unique) most parsimonious inversion scenario from each species to that ancestor. We note that the last common ancestor of human, mouse and rat should be on the evolutionary path between this median ancestor and human.

RepeatMasker annotations of the insertions of repetitive elements (Fig. 7). The rodent repeat database used by RepeatMasker was greatly expanded by analysing the rat and mouse genomes⁸⁹, but it is clear that not all repeats are being recognized, especially the older ones. Thus, these estimates of the amount of rodent repeats represent lower bounds.

About a billion nucleotides (39% of the euchromatic rat genome) align in all three species, constituting an 'ancestral core' that is retained in these genomes. This ancestral core contains 94–95% of the known coding exons and regulatory regions. Comparisons between the human and mouse genomes, using transposon relics retained in both species ('mammalian ancestral repeats') to model neutral evolution, have been used to estimate the fraction of the human genome that is accumulating substitutions more slowly than the neutral rate in both lineages since their divergence, and hence may be under some level of purifying selection³. Depending on details of methodology, such estimates have ranged between about 4% and 7%^{3,90,91}. The levels of three-way conservation observed here between the human, mouse and rat genomes in the ancestral core lend further support to these earlier estimates, giving values in the range of 5–6% when measured by two quite different methods (see Methods and ref. 92). In this constrained fraction, non-coding regions outnumber coding regions regardless of the strength of constraint⁹², an observation that supports recent comparative

analyses limited to subsets of the genome^{93,94}. The preponderance of non-coding elements in the most constrained fraction of the genome underscores the likelihood that they play critical roles in mammalian biology.

About 700 Mb (28%) of the rat euchromatic genome aligns only with the mouse. At least 40% of this comprises of rodent-specific repeats inserted on the branch from the primate–rodent ancestor to the murid ancestor, and some of the remainder can be recognized as mammalian ancestral repeats whose orthologues were deleted in the human lineage (Fig. 7). Another part is likely to consist of single-copy ancestral DNA deleted in the human lineage but retained in rodents. Although this 700 Mb of rodent-specific DNA is primarily neutral, it may also contain some functional elements lost in the human lineage in addition to sequences representing gains of rodent-specific functions, including some coding exons⁹⁵.

The remainder of the euchromatic rat genome (726 Mb, 29%) aligns with neither mouse nor human (Fig. 7). At least half of this (15% of the rat genome) consists of rat-specific repeats, and another large fraction (8% of the rat genome) consists of rodent-specific repeats whose orthologues are deleted in the mouse.

Substitution rates

The alignment data allow relatively precise estimates of the rates of neutral substitutions and microindel events (≤ 10 bp). Both synonymous fourfold degenerate ('4D') sites in protein-coding regions and sites in mammalian ancestral repeats were used in this analysis, as in previous studies comparing human and mouse^{3,96}. We additionally used a class of primarily neutral sites whose identification is made uniquely possible by the addition of the rat genome sequence: namely, the rodent-specific sites discussed above, identified by their failure to align to human sequence.

Our estimates for the neutral substitution level between the two rodents range from 0.15 to 0.20 substitutions per site, while estimates for the entire tree of human, mouse and rat range from 0.52 to 0.65 substitutions per site (Fig. 5). This difference was predictable because of the evolutionary closeness of the two rodents. For all classes of neutral sites analysed, however, the branch connecting the rat to the common rodent ancestor is 5–10% longer than the mouse branch (Fig. 5a). Thus, for as yet unknown reasons, the rat lineage has accumulated substantially more point substitutions than the mouse lineage since their last common ancestor.

We also analysed four-way alignments including sequence from orthologous ancestral repeats in human, mouse and rat, along with the repeat consensus sequences, which approximate the sequence of the progenitor of the corresponding repeat family (Methods). These alignments allow us to distinguish substitutions on the branch from the primate–rodent ancestor to the rodent ancestor from substitutions on the branch descending to human⁷⁷. This revealed an overall speed-up in rodent substitution rates relative to human of about three-to-one, larger than estimated previously³, but consistent with other more recent studies which also use multiple sequence alignments^{77,97,98}.

Estimates for rates of microdeletion events are, for all branches, approximately twofold higher than rates of microinsertion (Fig. 5b), suggesting a fundamental difference in the mechanisms that generate these mutations. Furthermore, there are substantial rate differences for each class of event between the various lineages. In particular, the rat lineage has accumulated microdeletions more rapidly than the mouse, while the opposite holds true for microinsertions. As with substitutions, both microinsertion and microdeletion rates are substantially slower in the human lineage. The size distribution of microindels (1–10 bp) on the rat branch was heavily weighted towards the smallest indels: 45% of indels are single bases, 18% are 2 bp, 10% are 3 bp, 8% are 4 bp, and so on, monotonically decreasing. Separate distributions for insertions and for deletions were similar, as were distributions of indel sizes on the mouse branch.

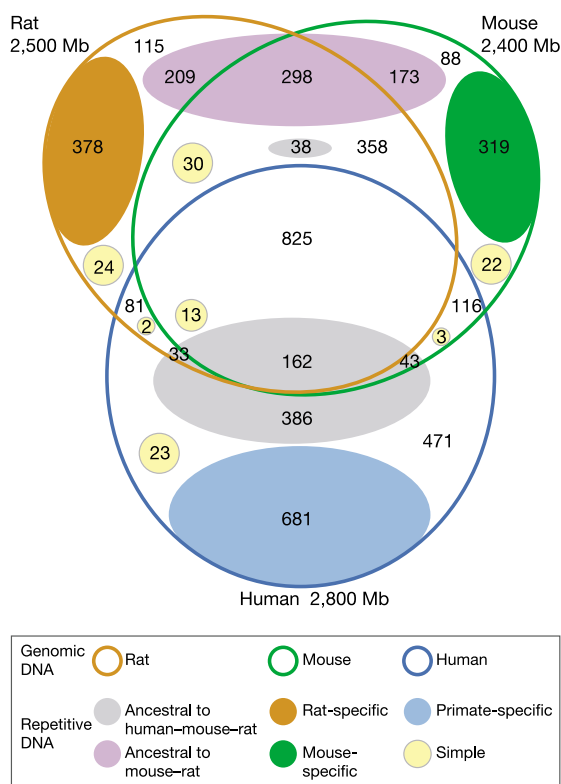


Figure 7 Aligning portions and origins of sequences in rat, mouse and human genomes. Each outlined ellipse is a genome, and the overlapping areas indicate the amount of sequence that aligns in all three species (rat, mouse and human) or in only two species. Non-overlapping regions represent sequence that does not align. Types of repeats classified by ancestry: those that predate the human–rodent divergence (grey), those that arose on the rodent lineage before the rat–mouse divergence (lavender), species-specific (orange for rat, green for mouse, blue for human) and simple (yellow), placed to illustrate the approximate amount of each type in each alignment category. Uncoloured areas are non-repetitive DNA—the bulk is assumed to be ancestral to the human–rodent divergence. Numbers of nucleotides (in Mb) are given for each sector (type of sequence and alignment category). Detailed results are tabulated (Supplementary Table SI-1).

Male mutation bias

As mouse and rat are similar in generation time and number of germline cell divisions^{99,100}, we investigated a potential sex bias in different types of observed genome changes. We compared substitution and indel rates between the X chromosome and autosomes in ancestral repeat sites (~5 Mb and ~100 Mb in total for X and autosomes, respectively¹⁰¹). We discovered that in rodents, small indels (<50 bp) are male-biased, with a male-to-female rate ratio of ~2.3. This is in contrast to a recent study in primates, based on a substantially smaller data set, that indicates no sex bias in small indels¹⁰². Our male-to-female nucleotide substitution rate ratio in rodents is ~1.9, confirming earlier reports^{103,104}. When substitution rates are compared for all sites aligned between mouse and rat (~78 Mb and ~1,691 Mb, respectively), we again observe an approximately twofold excess of small indels and nucleotide substitutions originating in males compared with females¹⁰¹. Interestingly, the ratio in the number of cell divisions between the male and female germlines is also about two^{99,100}, suggesting that these substitutions may arise from mutations that occur primarily during DNA replication.

G+C content and CpG islands

The G+C content of the rat varies significantly across the genome (Fig. 8a), and the distribution more closely resembles that of mouse than human. The variation in G+C content is coupled with differences in the distribution of CpG islands—short regions that are associated with the 5' ends of genes and gene regulation^{2,3,105}, and that escape the depletion of CpG dinucleotides that occurs from

deamination of methylated cytosine^{2,105}. The 2.6 Gb rat genome assembly (including unmapped sequences) contains 15,975 CpG islands in non-repetitive sequences of the genome. This is similar to the 15,500 CpG islands reported in the 2.5 Gb mouse genome³, but far fewer than the 27,000 reported in the human genome^{2,3,105}.

A summary of the CpG island distributions by chromosome is given in Fig. 8b. Chromosome X, with a low G+C content of 37.7%, has the fewest islands (362) and the lowest density of islands (2.6 per Mb). Chromosome 12 is at the other end of the range with a G+C content of 43.5% and the highest density of CpG islands (11.5 islands per Mb). This is similar to chromosome 10, with 11.3 islands per Mb. The average density of CpG islands is 5.7 islands per Mb over the whole genome and 5.9 CpG islands per Mb averaged by chromosome, which is similar to the distribution in mouse³. Neither rodent genome shows the extreme outliers in CpG island density that are seen for human chromosome 19 (ref. 2). The density of CpG islands in the rat genome correlates positively with the density of predicted genes (R of 0.96) (Fig. 8b).

These data show that the overall changes in CpG island content predate the rat–mouse split and are consistent with the accelerated loss of CpG dinucleotides in rodents compared with humans^{105,106}. It remains possible, however, that occurrences such as the greater number of human regions with extremely high G+C content are due to distributional changes mostly in the primate, rather than in the rodent lineage.

Shift in substitution spectra between mouse and rat

The non-repetitive fraction of the rat genome is enriched for G+C content relative to the mouse genome, by ~0.35% over 1.3 billion nucleotides. This is a subtle but substantial difference that may be explained, at least in part, by differences in the spectra of mutation events that have accumulated in the mouse and rat lineages. We analysed all alignment columns in which substitution events can be assigned to either the mouse or the rat lineage, by virtue of a nucleotide match between human and only one rodent⁹²; note that this is a small minority of substitutions. Of the ~117 million alignment columns meeting this criteria, ~60 million involve a change in the rat lineage versus ~57 million in the mouse, reflecting the increase in rates of point substitution in the rat lineage (Fig. 5b). While 50% of these changes in rat involve a substitution from an A/T to a G/C, these events constitute only 47% of all mouse changes. The complementary change, G/C to A/T, exhibits relative excess in the mouse versus the rat lineage (38% versus 35%, respectively). No substantial difference between changes that do not alter G+C content is observed. In addition, this bias is not confined to particular transition or transversion events, nor can it be explained simply as a result of divergent substitution rates of CpG dinucleotides (data not shown). Thus, this shift appears to be a general change that results in an increase in G+C content in the rat genome. Biochemical changes in repair or replication enzymes might be responsible, and the observation that recombination rates are slightly higher in rat than in mouse¹⁰⁷ may suggest a role for G+C-biased mismatch repair^{108,109}. However, population genetic factors, such as selection, cannot be ruled out.

Evolutionary hotspots

Comparison of the two rodent genomes, using human as outgroup, reveals regions that are conserved yet under different levels of constraint in mouse and rat. These regions may have distinct functional roles and contribute to species-specific differences. Analysis of the MAVID alignments¹¹⁰ revealed 5,055 regions ≥100 bp, in which there was at least a tenfold difference in the estimated number of substitutions per site on the mouse and rat branches. To avoid alignment problems and fast-evolving regions, the analysis was restricted to regions where the human branch had <0.25 substitutions per site¹¹¹. These regions are enriched twofold in transcribed regions: 39% of mouse hotspots were found in the

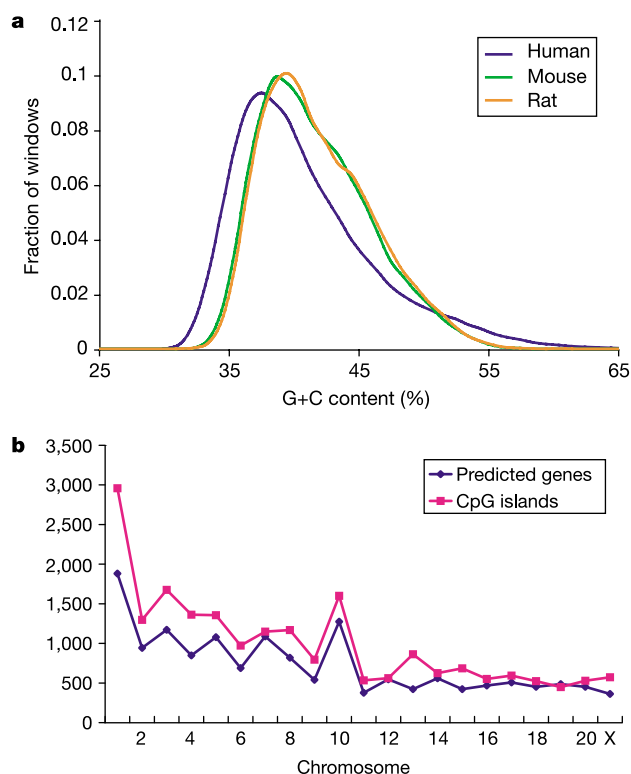


Figure 8 Base composition distribution analysis. **a**, The fraction of 20 kb non-overlapping windows³ with a given G+C content is shown for human, mouse and rat. **b**, The number of Ensembl-predicted genes per chromosome and the number of CpG islands per chromosome. The density of CpG islands averages 5.9 islands per Mb across chromosomes and 5.7 islands per Mb across the genome. Chromosome 1 has more CpG islands than other chromosomes, yet neither the island density nor ratio to predicted genes exceeds the normal distribution. The number of CpG islands per chromosome and the number of predicted genes are correlated ($R^2 = 0.96$).

18% of the mouse genome covered by RefSeq genes; and 17% of the rat hotspots were found in the 8% of the rat genome covered by RefSeq genes. Similar numbers are observed when examining coding exon and EST regions (not shown). Half of all hotspots in the mouse genome lie totally in non-coding regions. Many hotspots are several hundred bases long, with average length 190 ± 86 bp. Future work aimed at identifying the genomic differences that contribute to phenotypic evolution may benefit from analyses such as these, which will become more powerful as the repertoire of mammalian genome sequences expands.

Covariation of evolutionary and genomic features

To illustrate the genomic and evolutionary landscape of a single rat chromosome in depth, we characterized features for rat chromosome 10 at 1 Mb resolution (Fig. 9). This high-resolution analysis uncovered strong correlations between certain microevolutionary features^{89,92,98}. Particularly strongly correlated are the local rates of microdeletion ($R^2 = 0.71$; Fig. 9a), microinsertion ($R^2 = 0.56$; Fig. 9a), and point substitution ($R^2 = 0.86$; Fig. 9b) between the two independent lineages of mouse and rat. In addition, microinsertion rates are correlated with microdeletion rates ($R^2 = 0.55$; Fig. 9a). These strong correlations are also observed in an independent genome-wide analysis, both on the original data and after factoring out the effects of G+C content (not shown, see Supplementary Information).

Perhaps surprisingly, substantially less correlation is seen between microindel and point substitution rates (compare Fig. 9a and b). The amount of correlation varies among chromosomes (not

shown), but is generally weaker than the relationships mentioned above. Further studies will be required to determine whether local evolutionary pressures, which must have remained stable since the separation of the mouse and rat lineages, differentially drive microindel and point substitution rates.

We also find that the local point substitution rate in sites common to human, mouse and rat strongly correlates with that in rodent-specific sites ($R^2 = 0.57$; Fig. 9b, blue line versus red/green). These two classes of sites, while interdigitated at the level of tens to thousands of bases, constitute sites that are otherwise evolutionarily independent. This result confirms that local rate variation is not solely determined by stochastic effects and extends, at high resolution, the previously documented regional correlation in rate between 4D sites and ancestral repeat sites^{3,96}.

Evolution of genes

A substantial motivation for sequencing the rat genome was to study protein-coding genes. Besides being the first step in accurately defining the rat proteome, this fundamental data set yields insights into differences between the rat and other mammalian species with a complete genome sequence. Estimation of the rat gene content is possible because of relatively mature gene-prediction programs and rodent transcript data. Mouse and human genome sequences also allow characterization of mutational events in proteins such as amino acid repeats and codon insertions and deletions. The quality of the rat sequence also allows us to distinguish between functional genes and pseudogenes.

We estimate (on the basis of a subset) that 90% of rat genes

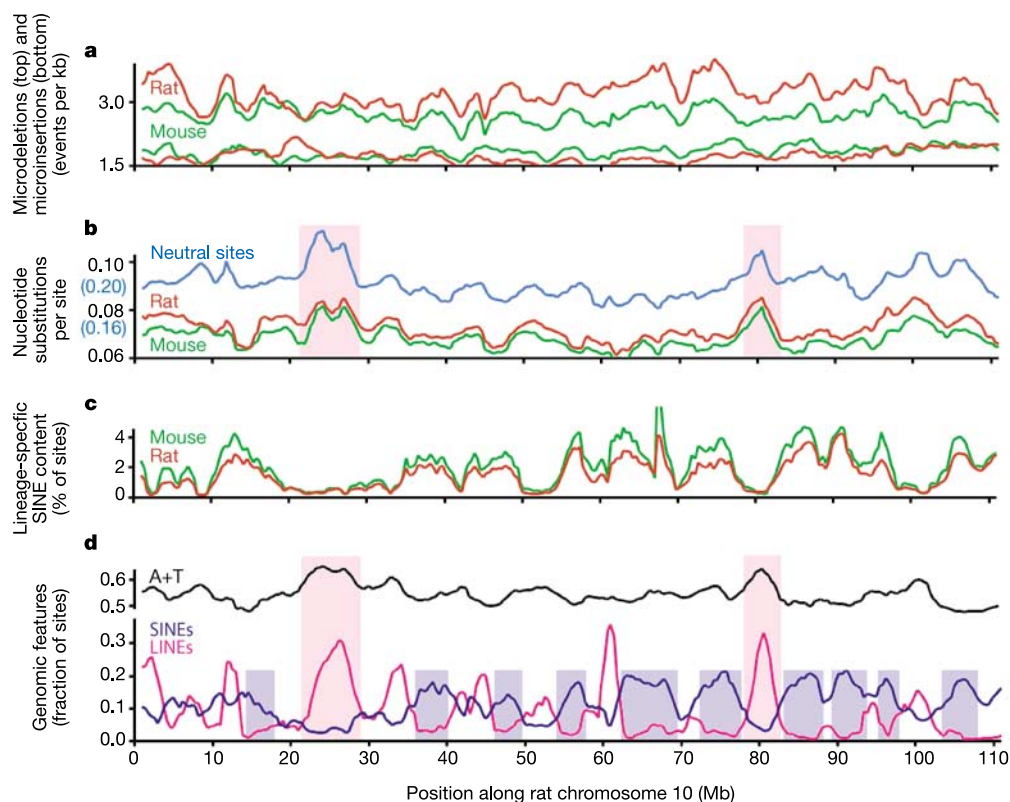


Figure 9 Variability of several evolutionary and genomic features along rat chromosome 10. **a**, Rates of microdeletion and microinsertion events (less than 11 bp) in the mouse and rat lineages since their last common ancestor, revealing regional correlations.

b, Rates of point substitution in the mouse and rat lineages. Red and green lines represent rates of substitution within each lineage estimated from sites common to human, mouse and rat. Blue represents the neutral distance separating the rodents, as estimated from rodent-specific sites. Note the regional correlation among all three plots, despite being

estimated in different lineages (mouse and rat) and from different sites (mammalian versus rodent-specific). **c**, Density of SINEs inserted independently into the rat or mouse genomes after their last common ancestor. **d**, A+T content of the rat, and density in the rat genome of LINEs and SINEs that originated since the last common ancestor of human, mouse and rat. Pink boxes highlight regions of the chromosome in which substitution rates, A+T content and LINE density are correlated. Blue boxes highlight regions in which SINE density is high but LINE density is low.

possess strict orthologues in both mouse and human genomes. Our studies also identified genes arising from recent duplication events occurring only in rat, and not in mouse or human. These genes contribute characteristic features of rat-specific biology, including aspects of reproduction, immunity and toxin metabolism. By contrast, almost all human 'disease genes' have rat orthologues. This emphasizes the importance of the rat as a model organism in experimental science.

Construction of gene set and determination of orthology

The Ensembl gene prediction pipeline¹¹² predicted 20,973 genes with 28,516 transcripts and 205,623 exons (Methods). These genes contain an average of 9.7 exons, with a median exon number of 6.0. At least 20% of the genes are alternatively spliced, with an average of 1.3 transcripts predicted per gene. Of the 17% single exon transcripts, 1,355 contain frameshifts relative to the predicted protein and 1,176 are probably processed pseudogenes. Of the 28,516 transcripts, 48% have both 5' and 3' untranslated regions (UTRs) predicted and 60% have at least one UTR predicted.

These gene predictions considered homology to other sequences, including 26,949 rodent proteins, 4,861 non-rodent, vertebrate proteins, 7,121 rat complementary DNAs from RefSeq and EMBL, and 31,545 mouse cDNAs from Riken, RefSeq and EMBL. The majority (61%) of transcripts are supported by rodent transcript evidence. When combined with additional private EST data, the fraction of genes supported by transcript evidence could be increased to 72%¹¹³.

A number of other *ab initio* (GENSCAN¹¹⁴, GENEID¹¹⁵), similarity-based (FGENESH++; ref. 116) and comparative (SGP¹¹⁷, SLAM¹¹⁸, TWINSKAN^{119–121}) gene-prediction programs were used to analyse the rat genome. The number of genes predicted by these programs ranged from 24,500 to 47,000, suggesting coding densities ranging from 1.2% to 2.2%. The coding fraction of RefSeq genes covered by these predictions ranged from 82% to 98%. Such comparative *ab initio* programs using the rat genome were successfully used to identify and experimentally verify genes missed by other methods in rat¹²¹ and human¹²². The predictions of these programs can be accessed through the UCSC genome browser and Ensembl websites.

RefSeq genes (20,091 human, 11,342 mouse and 4,488 rat) mapped onto genome assemblies with BLAT¹²³ and the UCSC browser revealed that the number of coding exons per gene and average exon length were similar in the three species. Differences were observed in intron length, with an average of 5,338 bp in human, 4,212 bp in mouse and 5,002 bp in rat. These differences were also found in a smaller collection of 6,352 confidently mapped orthologous intron triads (see 'Conservation of intronic splice signals' section below): average intron lengths in this collection were 4,240 bp in human, 3,565 bp in mouse and 3,638 bp in rat.

Properties of orthologous genes

Orthology relationships were predicted on the basis of BLASTp reciprocal best-hits between proteins of genome pairs (human–rat, rat–mouse and mouse–human)³ (Supplementary Information). Using these methods and the ENSEMBL prediction sets, 12,440

rat genes showed clear, unambiguous 1:1 correspondence with a gene in the mouse genome. This is an underestimate, because random sampling of different classes of rat genes with less stringent criteria for comparison to mouse always identified additional gene pairs. Errors arose from pseudogene misclassification, sequence loss, duplication or fragmentation in assemblies; and missing or inappropriate gene predictions, including coding-gene predictions from non-coding RNAs. Taking these errors into account, we estimate the true proportion of 1:1 orthologues in rat and mouse genomes to lie between 86 and 94% (Methods). The remaining genes were associated with lineage-specific gene family expansions or contractions. These overall observations are consistent with a careful analysis of rat proteases showing that 93% of these genes have 1:1 orthologues in mouse^{124,125}.

Surprisingly, a similar proportion (89 to 90%) of rat genes possessed a single orthologue in the human genome. Because human represents an outgroup to the two rodents, it was expected that mouse and rat would share a higher fraction of orthologues. A close inspection of gene relationships indicates that these findings may suffer from incompleteness of rodent genome sequences, together with problems of misassembly and gene prediction within clusters of gene paralogues.

Further analysis of orthologous pairs considered the occurrence of nucleotide changes within protein-coding regions that reflected synonymous or non-synonymous substitutions. The majority of these studies measured evolutionary rates by determination of K_A (number of non-synonymous substitutions per non-synonymous site) and K_S (number of synonymous substitutions per synonymous site). K_A/K_S ratios of less than 0.25 indicate purifying selection, values of 1 suggest neutral evolution, and values greater than 1 indicate positive selection¹²⁶.

Evolutionary rates were first calculated from a reduced set of orthologue pairs that are embedded in orthologous genomic segments and are related by conservative values of K_S (Table 3) (Methods). A slight increase in median K_S values for rat–human as compared with mouse–human, was found, indicating that the rat lineage has more neutral substitutions in gene coding regions than the mouse lineage. Sequence conservation values were similar to those previously found using smaller data sets^{127,128}, and the overall trend is consistent with results of other evolutionary rate analyses discussed above (Fig. 5).

Next, we investigated examples of rat genes shared with mouse, but with no counterparts in human. Such genes might be rapidly evolving so that homologues are not discernible in human, or they might have arisen from non-coding DNA, or their orthologues in the human lineage might have formed pseudogenes. Thirty-one Ensembl rat genes were collected that have no non-rodent homologues in current databases (Methods). These are twofold over-represented among genes in paralogous gene clusters, and threefold over-represented among genes whose proteins are likely to be secreted. This is consistent with observations³ that clusters of paralogous genes, and secreted proteins, evolve relatively rapidly. Detailed examination of the 31 genes using PSI-BLAST determined that ten genes cannot be assigned homology relationships to experimentally described mammalian genes. These ten rodent-

Table 3 One-to-one orthologous genes in human, mouse and rat genomes

	Human–mouse	Human–rat	Mouse–rat
1:1 orthologue relationships	11,084	10,066	11,503
Median K_S values*	0.56 (0.39–0.80)	0.57 (0.40–0.82)	0.19 (0.13–0.26)
Median K_A/K_S values*	0.10 (0.03–0.24)	0.09 (0.03–0.21)	0.11 (0.03–0.28)
Median % amino acid identity*	88.0% (74.4–96.3%)	88.3% (75.9–96.4%)	95.0%† (88.0–98.7%)
Median % nucleotide identity*	85.1% (77.4–90.0%)	85.1% (77.8–89.9%)	93.4% (89.2–95.7%)

Data obtained from Ensembl, *Homo sapiens* version 11.31 (24,841 genes), *Mus musculus* version 10.3 (22,345 genes), *Rattus norvegicus* version 11.2 (21,022 genes).

*Numbers in parentheses represent the 16th and 83rd percentiles.

†This value is consistent with previous findings (93.9% in ref. 130).

specific genes may have evolved particularly rapidly, or have non-coding DNA homologues, or be erroneous predictions.

The paucity of rodent-specific genes indicates that *de novo* invention of complete genes in rodents is rare. This is not unexpected, because the majority of eukaryotic protein-coding genes are modular structures containing coding and non-coding exons, splicing signals and regulatory sequences, and the chances of independent evolution and successful assembly of these elements into a functional gene are small, given the relatively short evolutionary time available since the mouse–rat split. However, individual rodent-specific exons may arise more frequently, particularly if the exon is alternatively spliced¹²⁹. Applying a K_A/K_S ratio test^{130,131} to sequences that align only between rat and mouse, we identified 2,302 potential novel rodent-specific exons, with EST support, in BLASTZ alignments of rat and mouse sequences. None of these individual exons matched human transcripts, but approximately half (1,116) appear to be present in alternative splice forms found in rodents. We speculate that these exons contain the few successful lineage-specific survivors of the constant process of gene evolution, by birth and death of individual exons.

Indels and repeats in protein-coding sequences

In contrast to small indels occurring in the bulk of the genome (above), indels within protein-coding regions are probably lethal, or deleterious and so are rapidly removed from the population by purifying selection. Indel rates within rat coding sequences were 50-fold lower than in bulk genomic DNA¹³². The whole genome excess of deletions compared with insertions (Fig. 5b) was also evident in coding sequences. The magnitude was less, with a genome-wide deletion-to-insertion ratio of 3.1:1 reducing to 1.7:1 in the rat. In mouse this value reduced from 2.5:1 to 1.1:1 (ref. 132). These data suggest that deletions are ~16% more likely than insertions to be removed from coding sequences by selection.

Owing to the triplet nature of the genetic code, indels of multiples of three nucleotides in length (3_n indels) are less likely to be deleterious. Direct comparison of 3_n indel rates between bulk DNA (0.77 indels per kb for mouse, 0.83 indels per kb for rat) and coding sequence (0.087 indels per kb for mouse and 0.084 indel per kb for rat) showed that 3_n indels were ninefold under-represented in coding sequences. At least 44% of indels were duplicative insertion or deletion of a tandemly duplicated sequence, collectively termed sequence slippage¹³². Sequence slippage contributed approximately equally to observed insertions and deletions. The overall excess of deletions could be attributed specifically to an excess of non-slippage deletion over non-slippage insertion in both mouse and rat lineages¹³². Of the slippage indels, 13% were in the context of trinucleotide repeats ($n > 2$, excluding the inserted or deleted sequence) which are known to be particularly prone to sequence slippage and encode homopolymeric amino acid tracts^{133,134}.

To gain better understanding of dynamic changes in the length of homopolymeric amino acid tracts on gene evolution and disease susceptibility, we searched for other characteristics of amino acid repeat variation by analysing all size-five or longer amino acid repeats in a data set of 7,039 rat, mouse and human orthologous protein sequences¹³⁵. Most species-specific amino acid repeats (80–90%) were found in indel regions, and regions encoding species-specific repeats were more likely to contain tandem trinucleotide repeats than those encoding conserved repeats. This was consistent with the involvement of slippage in the generation of novel repeats in proteins and extended previous observations for glutamine repeats in a more limited human–mouse data set¹³⁶.

The percentage of proteins containing amino acid repeats was 13.7% in rat, 14.9% in mouse and 17.6% in human¹³⁵. The most frequently occurring tandem amino acid repeats were glutamic acid, proline, alanine, leucine, serine, glycine, glutamine and lysine. Using the same threshold size cut-off, tandem trinucleotide repeats

were significantly more abundant in human than in rodent coding sequences, in striking contrast to the frequencies observed in bulk genomic sequences (29 trinucleotide repeats per Mb in rat, 32 repeats per Mb in mouse and 13 repeats per Mb in human, see discussion of the general simple repeat structure below). The conservation of human repeats was higher in mouse (52%) than in rat (46.5%), suggesting a higher rate of repeat loss in the rat lineage than the mouse lineage.

Functional consequences of these in-frame changes in rat, mouse and human were investigated¹³² through clustering of proteins based on annotation of function and cellular localization¹¹², and mapping indels onto protein structural and sequence features. The rate that indels accumulated in secreted (3.9×10^{-4} indels per amino acid) and nuclear (4.0×10^{-4}) proteins is approximately twice that of cytoplasmic (2.4×10^{-4}) and mitochondrial (1.4×10^{-4}) proteins. Likewise, ligand-binding proteins acquire indels (3.1×10^{-4}) at a higher rate than enzymes (2.1×10^{-4})¹³². These trends exactly mirror those observed for amino acid substitution rates³, suggesting tight coupling of selective constraints between indels and substitutions. Transcription regulators showed the highest rate of indels (4.3×10^{-4}), a finding that may relate to the over-representation of homopolymorphic amino acid tracts in these proteins¹³⁵.

Known protein domains exhibited 3.3-fold fewer indels than expected by chance, again paralleling nucleotide substitution rate differences between domains and non-domain sequences³. Of the protein-sequence and structural categories considered (transmembrane, protein domain, signal peptide, coiled coil and low complexity), the transmembrane regions were the most refractory to accumulating indels, exhibiting a sixfold reduction compared with that expected by chance. Low-complexity regions were 3.1-fold enriched, reflecting their relatively unstructured nature and enrichment in indel-prone trinucleotide repeats. Mapping of indels onto groups of known structures revealed that indels are 21% more likely to be tolerated in loop regions than the structural core of the protein¹³².

We observed that indel frequency and amino acid repeat occurrence both correlated positively with the G + C coding sequence content of the local sequence environment^{132,135}. This may be explained in part by the correlation of polymerase slippage-prone trinucleotide repeat sequences and G + C content¹³⁵. There is also a positive correlation between CpG dinucleotide frequency and coding sequence insertions, but not deletions. This effect diminishes rapidly with increasing distance from the site of the insertion¹³².

Transcription-associated substitution strand asymmetry

A recent study reported a significant strand asymmetry for neutral substitutions in transcribed regions¹³³. Within introns of nine genes, the higher rate of A→G substitutions over that of T→C substitutions, together with a smaller excess of G→A over C→T substitutions, leads to an excess of G+T over C+A on the coding strand (also verified on human chromosome 22). The authors¹³³ hypothesized that the asymmetries are a byproduct of transcription-

Table 4 **Strand asymmetry of substitutions in introns of rat genes**

Base frequencies on coding strand* (G+T)/(C+A)	Rat genome 1.060	
Ratio of purine transitions to pyrimidine transitions† Rate(A→G)/Rate(C→T)	Rat–mouse 1.036	Rat–human 1.036
Rate of transitions‡ Rate(A→G)/Rate(T→C)	Rat 1.058	Mouse 1.091
Rate(G→A)/Rate(C→T)	Rat 1.017	Mouse 1.00

*Computed from the rat genome.

†Computed from pairwise alignments.

‡Computed from three-way alignments.

coupled repair in germline cells. Examining the three-way alignments of rat, mouse and human, we verified that the strand asymmetries for neutral substitutions exist in introns across the genome (Table 4).

Under the assumption of independence of sequence positions, large sample normal approximations to the binomial distribution allow us to test whether the fraction of G+T exceeds 0.5, and whether the rate at the numerator exceeds the rate at the denominator for each of the ratios in Table 4. With the large amount of data provided by pooling introns genome-wide, the tests are all highly significant (P values $< 10^{-4}$), except for the rate of G→A in mouse, which does not significantly exceed that of C→T (P value = 0.6369). These asymmetries are also seen if the study is limited to ancestral repeat sites, excludes ancestral repeat sites, excludes CpG dinucleotides, is limited to positions flanked by sites that are identical in the aligned sequences (in the case of observations 2 and 3 in Table 4), or considers introns of RefSeq genes for human or mouse. Thus it appears that strand asymmetry of substitution events within transcribed regions of the genome is a robust genome-wide phenomenon.

Conservation of intronic splice signals

Using 6,352 human–mouse–rat orthologous introns from 976 genes (Methods), we examined the dynamics of evolution of consensus splice signals in mammalian genes. We found that intron class¹³⁷ is extremely well conserved: we did not observe any U2 to U12 intron conversion, or vice versa, nor within U12 introns did we find any switching between the major AT–AC and GT–AG subtypes, although such events are documented at larger evolutionary distances¹³⁷. In contrast, conversions between canonical GT–AG and non-canonical GC–AG subtypes of U2 introns are not uncommon. Only ~70% of GC–AG introns are conserved between human and mouse/rat, and only 90% are conserved between mouse and rat. Using human as the outgroup, we detected nine GT to GC conversions after divergence of mouse and rat (from 6,282 introns that were likely to have been GT–AG before human and rodents split), and two GC to GT conversions (from 34 GC–AG introns that probably predated the human and rodent split). These results give some indication of the degree to which mutation from T to C is tolerated in donor sites. The GC donor site appears to be better tolerated in introns with very strong donor sites, because in these introns the proportion of GC donor sites is ~11%, much higher than the 0.7% overall frequency of GC donor sites in U2 introns. Although we found a variety of other non-canonical configurations in U2 introns, very few are conserved, which suggests that most correspond to transient, evolutionarily unstable states, pseudogenes, or mis-annotations.

Gene duplications

Duplication of genomic segments represents a frequent and robust mechanism for generating new genes¹³⁸. Because there were no compelling data showing rat-specific genes arising directly from non-coding sequences, we examined gene duplications to measure their potential contribution to rat-specific biology. A previous study showed that gene clusters in mouse without counterparts in human are subject to rapid, adaptive evolution^{3,139}. We used two methods to identify recent gene duplications: methods that directly identified paralogous clusters, and methods that analysed genomic segmental duplications (see above).

Using the first approach, we found 784 rat paralogue clusters containing 3,089 genes (Methods). This was lower than in mouse (910 clusters/3,784 genes), but the difference probably reflects the larger number of gene predictions from the mouse assembly.

To investigate the timing of expansion of these individual families, we measured rates of local gene duplication and retention within clusters. BLAST is not suited to this^{140,141} and so we instead calculated the number of synonymous substitutions per

synonymous site (K_S) between all pairs of homologous genes; constructed K_S -derived phylogenetic trees; and predicted orthology or paralogy gene duplication events automatically from their topologies (Supplementary Information). The results showed that the neutral substitution rate varies among orthologues by approximately twofold (Fig. 10). This is similar to chromosomal variation shown previously by a study of mouse and human ancestral repeats³. Rates of change among ancestral gene duplications (those that predate the mouse–rat split) were relatively constant. Mouse-specific and rat-specific duplications occurred at similar rates, except for those with $K_S < 0.04$, which are reduced in mouse-specific duplications (Fig. 10). More data are required to determine whether this reduction is a biological effect, as it might be accounted for by different protocols for assembling mouse and rat genomes, which differentially collapse areas of nearly identical sequence.

The rat paralogue pairs that probably arose after the rat–mouse split (12–24 Myr ago) have K_S values of ≤ 0.2 (Table 3). We found 649 $K_S < 0.2$ gene duplication events in rat, a lower number than is found in mouse (755). For both rodents, this represents a likelihood of a gene duplicating of between 1.3×10^{-3} and 2.6×10^{-3} every Myr. These are necessarily estimates, because gene deletions, conversions and pseudogene formation are not considered. Interestingly, the data are consistent with a previous estimate for *Drosophila* genes, but are an order of magnitude lower than an estimate for *Caenorhabditis elegans* genes¹⁴⁰.

A subset of clusters have at least three gene duplications with $K_S < 0.2$ (Table 5). These are expected to be enriched in genes whose duplications persist as a consequence of positive selection. The group is dominated by genes involved in adaptive immune response and chemosensation⁸⁷. Inspection of the K_S -derived trees allowed us to infer the gene numbers in these clusters for the common ancestor of rat and mouse (that is, at $K_S = 0.2$), assuming no gene deletions or pseudogene generation (Table 5). Immunoglobulin, T-cell receptor α -chain, and α_{2u} -globulin genes appear to be duplicating at the fastest rates in the rat genome (Table 5). Since divergence with mouse, these rat clusters have increased gene content several-fold. This recapitulates previous observations that rapidly evolving and duplicating genes are over-represented in olfaction and odorant detection, antigen recognition and reproduction¹⁴².

An examination of duplicated genomic segments showed this enrichment for most of the same genes and also elements involved in foreign compound detoxification (cytochrome P450 and carboxylesterase genes)⁸⁷. Together, these are exciting findings because each of these categories can easily be associated with a

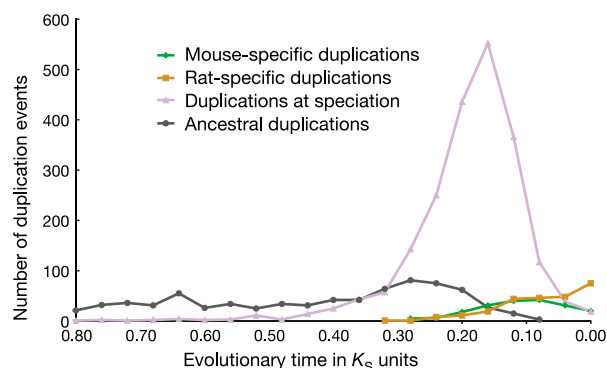


Figure 10 Variation in the frequency of gene duplications during the evolutionary histories of the rat and mouse. The sequence of gene duplication events was inferred from phylogenetic trees determined from pairwise estimates of genetic divergence under neutral selection (K_S , Methods). The median K_S value for mouse:rat 1:1 orthologues is 0.19. This value corresponds to the divergence time of mouse and rat lineages.

familiar feature of rat-specific biology, and further investigation could explain some differences between rats and their evolutionary neighbours.

Conservation of gene regulatory regions

As the third mammal to be fully sequenced, the rat can add significantly to the utility of nucleotide alignments for identifying conserved non-coding sequences^{143–147}. This power increases roughly as a function of the total amount of neutral substitution represented in the alignment^{97,98}, and rat adds about 15% to the human–mouse comparison (Fig. 5). Many conserved mammalian non-coding sequences are expected to have regulatory function, and can be predicted using further analyses based upon these alignments^{93,148–150}.

We applied such methods for detecting significantly conserved elements^{97,151} and scoring regulatory potential^{148,152} to the genome-wide human–mouse–rat alignments. Typical results show strong conservation for a coding exon, as well as for several non-coding regions (Fig. 11). For example, the intronic region in Fig. 11 contains 504 bp that are highly conserved in human, mouse and rat. The last 100 bp of this alignment block are identical in all three species. Peaks in regulatory potential score are correlated with conservation score, and in the highly conserved intronic segment, they are higher for the three-way regulatory potential score than for the two-way scores using human and just one rodent¹⁵². These data are illustrative, but form the foundation of ongoing efforts to identify genome sequences involved in gene regulation.

Requiring conservation among mammalian genomes greatly increases the specificity of predictions of transcription factor binding sites. Transcription factor databases such as TRANSFAC¹⁵³ contain known transcription factor binding sites and some knowledge of their distribution, but simply searching a sequence with these motifs provides little discriminatory power. For example, all of

the 85 known regulatory elements¹⁴⁸ and 151 functional promoters¹⁵⁴ have TRANSFAC matches, but so do 99% of the 2,049,195 mammalian ancestral repeats, most representing false-positive predictions. The introduction of conservation as a criterion for regulatory element identification greatly increases specificity, with only a modest cost in sensitivity. If we insist that the TRANSFAC matches be present and orthologously aligned in all three species—human, mouse and rat—then only 268 matches are recorded in ancestral repeats (0.01%), while 63 (74%) of the above matches in known regulatory elements and 121 (80%) in functional promoters are retained. Overall, using a set of 164 weight matrices for 109 transcription factors extracted from TRANSFAC¹⁵³, we find 186,792,933 matches in the April 2003 reference human genome sequence, but this was reduced to only 4,188,229 by demanding conservation in the human–mouse–rat three-way alignments. This is a 44-fold increase in specificity.

We examined one region in more detail: a complex *cis*-regulatory region consisting of a 4,000 bp segment containing two regulatory modules, hypersensitive sites 2 and 3 from the locus control region of the HBB complex^{155–157}. Considerable experimental work has identified six functional binding sites for the transcription factor GATA-1 in this segment. Requiring that matches to GATA-1 binding sites be conserved in all three species and occur within regions of strong regulatory potential is sufficient to find these six functional binding sites, and only these six, in the 4,000 bp segment. Thus, in this example we observed complete sensitivity and specificity by requiring this level of conservation.

Pseudogenes and gene loss

To complement the identification and analysis of protein-coding regions, we sought to examine rat pseudogenes. Using a previously described method^{158,159}, we found 18,755 pseudogenes in intergenic regions. Pseudogenes are normally not subjected to selective con-

Table 5 Recent gene duplications ($K_S < 0.2$) in the rat lineage

Cluster ID	Recent duplication events	Numbers of genes involved	Extant cluster size	Ancestral cluster size	Chromosome	Annotation	Process
249	38	53	60	22	4	Immunoglobulin κ -chain V	Immunity
640	38	47	53	15	15	TCR α -chain V	Immunity
346	25	35	44	15	6	Immunoglobulin heavy chain V	Immunity
190	22	42	168	146	3	Olfactory receptor	Chemosenation
578	16	28	59	43	13	Olfactory receptor	Chemosenation
400	15	26	82	67	8	Olfactory receptor	Chemosenation
743	15	21	37	22	20	Olfactory receptor	Chemosenation
72	12	22	102	90	1	Olfactory receptor	Chemosenation
500	12	18	32	20	10	Olfactory receptor	Chemosenation
51	6	7	16	10	1	Glandular kallikrein	Reproduction?
256	6	8	10	4	4	Vomeranase receptor V1R	Chemosenation
488	6	10	11	5	10	Olfactory receptor	Chemosenation
644	6	10	14	8	15	Granzyme serine protease	Immunity
4	5	6	9	4	1	Trace amine receptor, GPCR	Neuropeptide receptors?
248	5	9	15	10	4	Vomeranase receptor V1R	Chemosenation
393	5	10	31	26	8	Olfactory receptor	Chemosenation
522	5	8	19	14	10	Keratin-associated protein	Epithelial cell function
550	5	8	17	12	11	Olfactory receptor	Chemosenation
635	5	9	20	15	15	Olfactory receptor	Chemosenation
79	4	8	38	34	1	Olfactory receptor	Chemosenation
88	4	6	11	7	1	Olfactory receptor	Chemosenation
109	4	7	43	39	1	Olfactory receptor	Chemosenation
294	4	5	5	1	5	α_{2u} -globulin	Chemosenation
310	4	5	11	7	5	Olfactory receptor	Chemosenation
353	4	7	13	9	7	Olfactory receptor	Chemosenation
399	4	5	6	2	8	Ly6-like urinary protein	Chemosenation?
638	4	6	6	2	15	RNase A	Immunity
690	4	6	21	17	17	Prolactin paralogue	Reproduction
239	3	6	6	3	4	Prolactin-induced protein	Reproduction
253	3	4	5	2	4	Camello-like N-acetyltransferase	Developmental regulator
274	3	6	20	17	4	Ly-49 lectin natural killer cell protein	Immunity
297	3	4	5	2	5	Interferon- α	Immunity
523	3	4	6	3	10	Keratin-associated protein	Epithelial cell function
746	3	5	6	3	20	MHC class 1b (M10)	Chemosenation

Duplications involving retroviral genes, fragmented genes with internal repeats, and likely pseudogene clusters were removed from this list. Only gene clusters exhibiting at least three duplications are shown.

straint and therefore accumulate sequence modifications neutrally. Indeed, nearly all of our identified pseudogenes ($97 \pm 3\%$) evolved under neutrality according to a K_A/K_S test, and therefore are consistent with being pseudogenic.

We classified these pseudogenes according to whether they arose from retrotransposition, in which case they integrated into the genome randomly, or whether they arose from tandem duplication and neutral sequence substitution. Using human–rat synteny, we found that 80% of pseudogenes exhibited no significant similarity to the corresponding human orthologous region, and therefore were considered retrotransposed, processed pseudogenes. The total pseudogene count, and processed pseudogene proportion, are consistent with those found for human^{158,159}. These numbers are greater than those previously reported for mouse^{3,4}. However, reanalysis using the method employed here detects a similar pseudogene number (20,000) to that found for human and rat. This suggests that the rate of pseudogene creation is similar among these mammals.

As with the human genome^{159,160}, the largest group of rat pseudogenes (totalling 2,188), according to InterPro¹⁶¹, consists of ribosomal protein genes. Other large rat pseudogene families arose from olfactory receptors (552, see below), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (251), protein kinases (177), and RNA binding RNP-1 proteins (174). Pseudogenes homologous to a meiotic spindle-associated protein—spindlin¹⁶²—are particularly numerous in rat (at least 53 copies) compared with mouse (approximately three copies). This suggests that spindlin pseudogenes may have distributed rapidly by a recently active transposable element.

We investigated the much-studied metabolic enzyme GAPDH^{3,163}, and observed that: (1) the GAPDS gene arose from a duplication of the GAPDH gene; (2) biogenesis of the GAPDH pseudogenes has been occurring steadily over time both before and

after rodent–human and mouse–rat divergence; and (3) the GAPDS gene has undergone little retrotransposition in all three genomes compared with its relative, the GAPDH gene (consistent with respective gene-expression levels in the germ line).

In situ loss of rat genes

As an organism evolves, its need for certain genes may be reduced, or lost, owing to changes in its ecological niche. Loss of selective constraints leads to accumulation of nonsense and/or frameshift mutations without retrotransposition or duplication. These non-processed pseudogenes are interesting because they link environmental changes to genomic mutation events. However, predicted pseudogenes with disrupted reading frames might also be indicative of errors in genome sequence or assembly. By constraining the search to orthologous genomic regions, we identified 14 rat putative non-processed pseudogenes (Table 6) with apparently functional, single human and mouse orthologues. Half of these contain one in-frame stop or frameshift, whereas the remainder contain more. We expect this number of identified pseudogenic orthologues to be conservative because the methods employed required high fidelity of both gene prediction and orthologue identification in all three species (Methods).

Nevertheless, as only 14 recently evolved pseudogene candidates were identified, this indicates that the genome sequence and assembly (Rnor3.1) is of high quality. The improved quality of the most recent assembly is underscored by 11 additional candidate pseudogenes, predicted from rat assembly Rnor2.1, that are apparently functional, full-length genes in Rnor3.1. Consequently, some of the current 14 candidates, in particular those that are involved in fundamental processes of eukaryotic biology, may yet be ‘repaired’ by sequence changes in future assemblies, and thus be recognized as genic. However, genes associated with innate immunity (which is particularly susceptible to change via adaptive evolution), such as Forssman glycolipid synthetase and complement factor I, may yet be found to survive as true pseudogenes in the rat.

Non-coding RNA genes

We investigated the abundance and distribution of non-coding (nc)RNAs in rat. Cytoplasmic transfer (t)RNA gene identification in rodents is complicated by tRNA-derived identifier (ID) short interspersed nucleotide (SINEs) (B2 and ID). tRNAscan-SE predicted 175,943 tRNAs (genes and pseudogenes); however, the majority (175,285) were SINEs identified by RepeatMasker. This is far greater than the number found in mouse (24,402/25,078) or human (25/636). Of the remaining 666 predictions, 163 were annotated as tRNA pseudogenes and four were annotated as undetermined by tRNAscan-SE. An additional 68 predictions were removed because their best database match in either human, mouse or rat tRNA databases matched tRNAs with either a different amino acid or anticodon (violating the wobble rules that specify the distinct anticodons expected). The total of 431 tRNAs (including a single selenocysteine tRNA) identified in the rat genome is comparable to that for mouse—435 tRNAs (version mm2 from the UCSC genome browser)—and human—492 tRNAs (from the genomic tRNA database, <http://rna.wustl.edu/GtRDB/Hs/>). These three species share a core set of approximately 300 tRNAs, using a cutoff of $\geq 95\%$ sequence identity and $\geq 95\%$ sequence length.

A total of 454 ncRNAs (other than tRNAs) were identified by sequence comparison to known ncRNAs (Supplementary Information). These include 113 micro- (mi)RNAs, five ribosomal RNAs, 287 small nucleolar (sno)RNAs and small nuclear (sn)RNAs, 49 various other ncRNAs such as signal recognition particle (SRP) RNA, 7SK RNA, telomerase RNA, RNase P RNA, brain-specific repetitive (bsr)RNA, non-coding transcript abundantly expressed in brain (ntab)RNA, small cytoplasmic (sc)RNA and 626 pseudogenes. Complete 18S and 28S rRNA genes and more rRNAs were not identified, presumably owing to assembly issues.

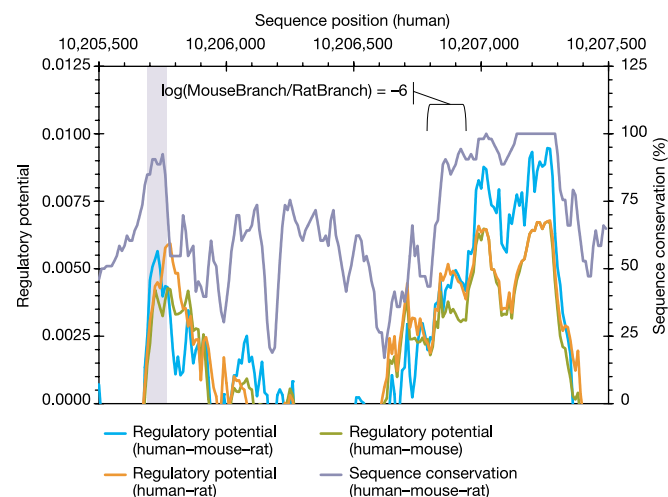


Figure 11 Close-up of PEX14 (peroxisomal membrane protein) locus on human chromosome 1 (with homologous mouse chromosome 4 and rat chromosome 5). Conservation score computed on three-way human–mouse–rat alignments (parsimony P values¹⁵¹) presents a clear coding exon peak (grey bar) and very high values in a 504 bp non-coding, intronic segment (right; last 100 bp of alignment are identical in all three organisms). The latter segment showed a striking difference between the inferred mouse and rat branch lengths^{110,111,222}; the grey bracket corresponds to a phylogenetic tree where the logarithm of mouse to rat branch-length ratio is -6 . Regulatory potential scores^{148,152} that discriminate between conserved regulatory elements and neutrally evolving DNA are calculated from three-way (human–mouse–rat) and two-way (human–rodent) alignments. Here the three-way regulatory potential scores are enhanced over the two-way scores.

Table 6 Candidate rat pseudogenes, orthologous to mouse and human functional genes

Mouse gene	Human gene	Strand	Rat genome coordinates*	Frameshifts/stops†	Annotation
ENSMUSG00000013611	ENSG00000174226	+	7:92752590–92807556	1/0	Sorting nexin
ENSMUSG00000024364	ENSG00000158402	+	18:62742414–62770427	2/0	Dual-specificity phosphatase CDC25c
ENSMUSG00000026293	ENSG00000077044	+	9:95634847–95692601	1/0	Diacylglycerol kinase δ
ENSMUSG00000026785	ENSG00000160447	+	3:9210762–9229984	5/0	Protein kinase PKN β
ENSMUSG00000026829	ENSG00000148288	+	3:7662414–7664521	2/2	Forssman glycolipid synthetase
ENSMUSG00000027426	ENSG00000125846	+	3:125918806–125924149	1/1	Zinc finger protein 133
ENSMUSG00000028000	ENSG00000138799	–	2:221272797–221304350	1/0	Complement factor I
ENSMUSG00000029203	ENSG00000078140	–	14:44385206–44441888	1/0	Ubiquitin-protein ligase E2 (HIP2)
ENSMUSG00000030270	ENSG00000144550	–	20:8332585–8362331	3/0	Copine (membrane trafficking)
ENSMUSG00000035449	ENSG00000167646	+	1:67374986–67381472	1/0	Cardiac troponin I
ENSMUSG00000037029	ENSG00000105261	–	1:82728049–82730272	1/0	Zinc finger protein 146
ENSMUSG00000037432	ENSG00000158142	+	9:42465695–42498651	1/1	Dysferlin-like protein
ENSMUSG00000039660	ENSG00000167137	–	3:9320401–9326997	4/0	Similar to yeast YMR310c RNA-binding protein
ENSMUSG00000042653	ENSG00000137634	+	8:49938446–49939091	1/0	Brush border 61.9 kDa-like protein

*Coordinates from rat v2.0.
†Mouse genes were used as templates for predicting rat pseudogenes.

Evolution of transposable elements

Most interspersed repeats are immobilized copies of transposable elements that have accrued substitutions in proportion to their time spent fixed in the genome (for introduction^{2,3,164–167}). About 40% of the rat genome draft is identified as interspersed repetitive DNA derived from transposable elements, similar to that for the mouse³ (Table 7) and lower than for the human (almost 50%²). The latter difference is mainly due to the lower substitution rate in the human lineage, which allows us to recognize much older (Mesozoic) sequences as interspersed repeats. Almost all repeats are derived from retroposons, elements that procreate via reverse transcription of their transcripts. As in mouse, there is no evidence for activity of DNA transposons since the rat–mouse split. Many aspects of the rat and the mouse genomes’ repeat structure are shared; here we focus on the differences.

LINE-1 activity in the rat lineage

The long interspersed nucleotide element (LINE)-1 (L1) is an autonomous retroelement, containing an internal RNA polymerase II promoter and two open reading frames (ORFs). The ORF1 product is an RNA binding protein with chaperone-like activity, suggesting a role in mediating nucleic acid strand transfer steps

during L1 reverse transcription¹⁶⁸, whereas ORF2 encodes a protein with both reverse transcriptase and DNA endonuclease activity. LINEs are characteristically 5’ truncated so that only a small subset extends to include the promoter region and can function as a source for more copies.

Many classes of LINE-like elements exist, but only L1 has been active in rodents. Over half a million copies, in variable stages of decay, comprise 22% of the rat genome. Although 10% of the human genome is comprised of L1 copies introduced before the rodent–primate split, owing to the fast substitution rate in the rodent lineage only 2% of the rat genome could be recognized as such. Thus, probably well over one-quarter of all rat DNA is derived directly from the L1 gene.

Following the mouse–rat split, L1 activity appears to have increased in rat. The 3’ UTR sequences defined six rat-specific L1 subfamilies, represented by 150,000 copies that cover 12% of the rat genome. L1 copies accumulated over the same period in mouse cover only 10% of the genome (Table 7). This higher accumulation of L1 copies could explain some of the size difference of the rat and mouse genome.

In addition to the traditional L1 elements, there are 7,500 copies

Table 7 Composition of interspersed repeats in the rat genome

	Rat				Mouse	
	Copies (× 10 ³)	Total length (Mb)	Fraction of genome (%)	Lineage-specific (%)	Fraction of genome (%)	Lineage-specific (%)
LINES	657	594.0	23.11	11.70	20.10	9.74
LINE-1	597	584.2	22.73	11.70	19.65	9.74
LINE-2	48	8.4	0.33	–	0.38	–
L3/CR1	11	1.4	0.06	–	0.06	–
SINES	1,360	181.3	7.05	1.52	7.78	1.80
B1(Alu)	384	42.3	1.65	0.16	2.53	0.92
B4(ID_B1)	359	55.4	2.15	0.00	2.25	0.00
ID	225	19.6	0.76	0.54	0.20	0.00
B2	328	55.2	2.15	0.68	2.29	0.74
MIR	109	13.0	0.51	–	0.56	–
LTR elements	556	232.4	9.04	1.84	10.28	2.85
ERV_class I	40	24.9	0.97	0.56	0.79	0.36
ERV_class II	141	83.4	3.24	1.02	4.13	1.73
ERV_L (III)	74	21.6	0.84	0.04	1.08	0.23
MaLRs	302	102.5	3.99	0.22	4.27	0.53
DNA elements	108	20.9	0.81	–	0.86	–
Charlie(hAT)	80	14.8	0.58	–	0.60	–
Tigger(Tc1)	18	4.0	0.16	–	0.17	–
Unclassified	14	7.3	0.28	–	0.37	–
Total	2,690	1,036	40.31	14.90	39.45	14.26
Small RNAs	8	0.6	0.03	0.01	0.03	0.01
Satellites	14	6.4	0.25	?	0.31	?
Simple repeats	897	61.1	2.38	?	2.41	?

Data for Rnor3.1 and October 2003 mouse (MM4), excluding Y chromosome, using the 17 December 2003 version of RepeatMasker. To highlight the differences between rat and mouse repeat content, columns 5 and 7 show the fractions of the genomes comprising lineage-specific repeats. The LINE-1 numbers include all HAL 1 copies, whereas all BC1 scRNA and >10% diverged tRNA-Ala matches, far more common than other small RNA pseudogenes and closely related to ID, have been counted as ID matches.

(10 Mb) of a non-autonomous element that is derived from L1 by deletion of most of its ORF2. A similar element, active in Mesozoic times, has been called HAL1 (for Half-a-LINE)¹⁶⁴. Given their low divergence, we conclude that the currently identified HAL1-like elements operated only a few million years ago in the mouse lineage (MusHAL1) and still propagate in the rat genome (RNHAL1). RNHAL1 contains only an ORF1, whereas MusHAL1 encoded an endonuclease as well, although no reverse transcriptase. The 5' 2,600 bases of RNHAL1 are 98% identical to the currently active L1 in rat (L1_Rn or L1mlvi2¹⁶⁹). Unlike ancient HAL1 elements, which shared the 3' UTR with a contemporary L1, the 3' end of RNHAL1 is unrelated to other repeats. The repeated origin and high copy number of HAL1s suggest that the ORF1 product, which binds strongly to its messenger RNA¹⁶⁸, may render this transcript a superior target for L1-mediated reverse transcription. In this way HAL1 resembles the non-autonomous, endogenous retrovirus-derived MaLR elements (below), which, for over 100 million years, retained only the retroviral gag ORF that encodes an RNA binding protein. A potential advantage of HAL1 over L1 is its shorter length, which, considering the usual 5' truncation of copies, increases the chance that a copy may include the internal promoter elements and become a source gene.

Different activity of SINEs in the rat and mouse lineage

The most successful usurpers of the L1 retrotransposition machinery, however, are SINEs. These are small RNA-derived sequences with an internal RNA polymerase III promoter. Recently, the human Alu SINE has been experimentally proven to be transposed by L1¹⁷⁰. Most SINEs share the 3' end with their associated LINE elements, like the Mesozoic mammalian LINE-2 (L2) and MIR pair, increasing the efficiency with which a LINE reverse transcriptase recognizes the 3' end of a dependent SINE. However, L1 does not show sequence specificity and rodent and primate SINE sequences are unrelated to L1. Although any transcript can be retroposed, as can be seen from the numerous processed pseudogenes in mammalian genomes, L1-dependent SINEs probably have features that make them especially efficient targets of the L1 reverse transcriptase.

Although before the radiation of most mammalian orders L1 was at least as active as L2, the L2-dependent MIR was the only known (and very abundant) SINE of that time. All of the currently active SINEs in different mammalian orders appear to have arisen after the demise of L2 (and consequently MIR), as though an opportunity (or necessity) arose for the creation and expansion of other SINEs.

Four different SINEs are distinguished in rat and mouse. The B1 element seems to share its origin from a 7SL RNA gene with the primate Alu¹⁷¹. This probably happened just before the rodent-primate split and after the speciation from most other eutherians, where Alu/B1 elements are not known. The other SINEs are rodent-specific and have tRNA-like internal promoter regions. ID elements consist only of this tRNA-like region, which in older ID copies closely match an Ala-tRNA from which it may have been derived. B4 resembles a fusion of an ID and B1 SINE. Finally, B2 has a tRNA-like region of unknown affiliation followed by a unique 120 bp region.

The fortunes of these SINEs during mouse and rat evolution have been different (Fig. 12). B4 probably became extinct before the mouse-rat speciation, while B2 has remained productive in both lineages, scattering >100,000 copies in each genome after this time. Interestingly, the fate of the B1 and ID SINEs has been opposite in rat and mouse. While B1 is still active in mouse, having left over 200,000 mouse-specific copies in its trail, the youngest of the 40,000 rat-specific B1 copies are 6–7% diverged from their source, indicating a relatively early extinction in the rat lineage. On the other hand, after the mouse-rat split only a few hundred ID copies may have inserted in mouse, whereas this previously minor SINE (~60,000 copies predate the speciation) increased its activity in rat to produce 160,000 ID copies.

Co-localization of SINEs in rat and mouse

Despite the different fates of SINE families, the number of SINEs inserted after speciation in each lineage is remarkably similar: ~300,000 copies. Reminiscent of the replacement of MIR by L1 driven SINEs, it seems that the demise of B1 in rat allowed the expansion of IDs. Moreover, these independently inserted and unrelated SINEs (ID and B1 share only a mechanism of retro-position) accumulated at orthologous sites: the density of rat-specific SINEs in 14,243 ~100 kb windows in the rat genome is highly correlated ($R^2 = 0.83$) with the density of mouse-specific SINEs in orthologous regions in mouse. To avoid including elements fixed before the speciation, only SINEs labelled lineage-specific on the basis of subfamily assignment (Methods⁸⁹) were tallied with a divergence from the consensus that was well below the 9% average for neutral sites (Fig. 5). These data corroborate and refine the observation of a strong correlation between the location of primate- and rodent-specific SINEs in 1 Mb windows³. At 100 kb, no correlation is seen for interspersed repeats other than SINEs.

Insertions of SINEs at the same location in different species have been reported^{172–174}, and the correlation could reflect the existence of conserved hotspots for SINE insertions. However, only five of ~800 human specific Alu elements have an Alu inserted within 100–200 bp in any of six other primate lineages^{174–176}. Likewise, gene conversions of shared Alus into lineage-specific copies were observed five times in the same set, too low a level to contribute significantly to the observed correlation^{174–176}.

Figure 9c displays the lineage-specific SINE densities on rat chromosome 10 and in the mouse orthologous blocks, showing a stronger correlation than any other feature. The cause of the unusual distribution patterns of SINEs, accumulating in gene-rich regions where other interspersed repeats are scarce, is apparently a conserved feature, independent of the primary sequence of the SINE and effective over regions smaller than isochores.

In the human genome, the most recent (unfixed) Alus are distributed similarly to L1, whereas older copies gradually take on the opposite distribution of SINEs^{2,164}. This suggested that SINEs insert in the same places as LINES, and that the typical SINE pattern is due to selection (or deletion bias) rather than a mechanistic insertion bias shared by all (unrelated) SINEs, but not by LINES that use the same insertion process. This led to a proposal that SINEs are preferentially maintained in regions where they can easily be expressed^{2,164}; if so, this could be the local feature conserved between mammalian genomes that leads to the strong correlation of local SINE densities in different mammals. However, we did not observe this temporal shift in SINE distribution pattern in mouse, nor currently in the rat genome, despite a considerable effort to define the potentially unfixed SINEs in both species (see ref. 89 for details). The observations in human could reflect a recent change in Alu behaviour, which would necessitate another explanation for the contrary insertion-preference of older Alus and all other SINEs.

Some regions of high LINE content coincide with regions that exhibit both higher AT content and an increased rate of point substitution (Fig. 9, pink rectangles). In a genome-wide analysis, LINE content correlates strongly with substitution rates, and about 80% of this correlation is explained by higher rates in AT-rich regions⁸⁹. SINE density shows the opposite correlation both on chromosome 10 (Fig. 9) and genome-wide⁸⁹.

These phenomena, in conjunction with an overall trend in substitution rates towards AT-richness, suggest a model in which quickly evolving regions accumulate a higher-than-average AT content, which attracts LINE elements. Although distinct cause-effect relationships such as this remain largely speculative, these results reinforce the idea that local genomic context strongly shapes local genomic features and rates of evolution.

Endogenous retroviruses and derivatives

The other major contributors to interspersed repeats in the rodent

genome are retrovirus-like elements. These have several 100 bp long terminal repeats (LTRs) with transcriptional regulatory sequences that flank an internal sequence that, in autonomous elements, encodes all proteins necessary for retrotransposition. All mammalian LTR elements are endogenous retroviruses (ERVs) or their non-autonomous derivatives. They fall into three groups, of which representatives in mouse are: murine leukaemia virus (MuLV) (class I), intracisternal A-particle (IAP) and MMTV (class II), and MERVL (class III).

The most productive retrovirus in mammals has been the class III element ERV-L, primarily through its ancient non-autonomous derivatives, called MaLRs, with 350,000 copies occupying ~5% of the rat genome (Table 7). Human ERV-L and MaLR copies are >6% diverged from their reconstructed source genes and must have died out around the time of human speciation from New World monkeys. In mouse, several thousand almost identical MaLR and ERV-L copies suggest sustained activity^{177–179}. In contrast, rat ERV-L activity must have been silenced a few million years ago, given that the least diverged MaLR and ERV-L (MTB_Rn and MT2_Rat1) copies differ by >4% from each other. Other class III ERVs were active earlier in rodent evolution, before the mouse–rat speciation.

In contrast to class III ERVs, class I and class II elements still thrive in rat. We reconstructed four rat-specific autonomous class I ERVs, of which two appear still active, and nine class II ERVs, of which four may still be active. The non-autonomous NICER and RAL elements represent over 60% of all rat-specific class I elements. The autonomous drivers of this group, RNNICER2 and 3, with several intact copies, are closely related to the mouse-specific MuLV. Among the potentially active autonomous class II ERVs are MYSERV_Rn, related to the Mys element in *Peromyscus*, and several IAP elements, one with a full-length envelope gene. The most prolific, still-active class II ERV, RNERVK3, is distantly related to the simian retroviruses and, like ERV-L and NICER, has spawned abundant non-autonomous elements characterized by closely related LTRs.

Simple repeats

Whereas the above interspersed repeats derive from transposed sequences, mammalian genomes also contain interspersed simple sequence repeats (SSRs), regions of tandemly repeated short (1–6 bp) units that probably arise from slippage during DNA replication and can expand and compress by unequal crossing

over. Remarkable differences were noted between the SSR contents of the human and mouse genomes³. Three times as many base pairs are contained in near (>90%) perfect SSRs in mouse than in human, and a 4–5-fold excess was revealed when excluding SSRs contained in or seeded by interspersed repeats (primarily SSRs derived from the poly A or simple repeat tails of SINEs and LINEs). SSRs are both more frequent and on average longer in mouse. Polypurine (or polypyrimidine) repeats are especially (tenfold) over-represented in the mouse genome. As discussed above, this contrasts sharply with the greater frequency of triplet repeats coding for amino acids in human than in the rodents.

Rat and mouse SSR contents show, perhaps not surprisingly, much smaller differences. They represent almost the same amount of the rat and mouse genomes (for >90% perfect elements, ~1.4% compared with 0.45% in human) and are of similar average length; for example, the average >90% perfect (CA)_n repeat, the most common SSR in mammals, is 42 bp long in mouse and 44 bp in rat. Some potentially significant differences are that polypurine SSRs are of similar average length but are 1.2-fold more common in mouse, whereas the rare SSRs containing CG dimers are 1.5-fold more frequently observed in rat.

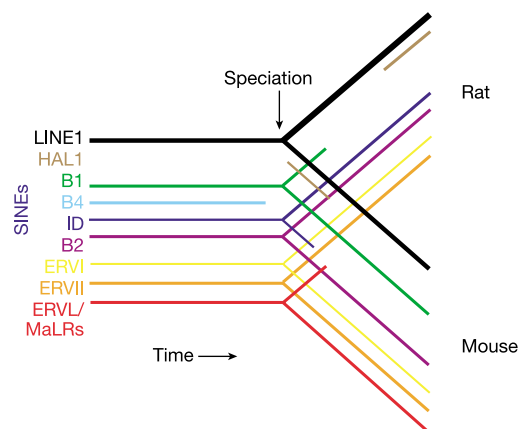


Figure 12 Historical view of rodent repeated sequences. Relationships of the major families of interspersed repeats (Table 7) are shown for the rat and mouse genomes, indicating losses and gains of repeat families after speciation. The lines indicate activity as a function of time. Note that HAL1-like elements appear to have arisen in both the mouse and rat lineages.

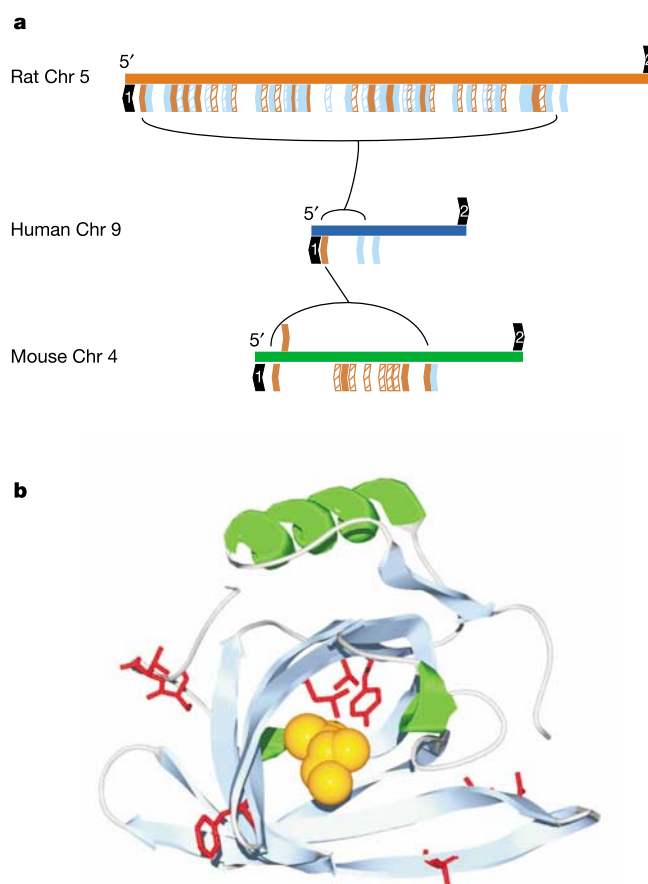


Figure 13 Adaptive remodelling of genomes and genes. **a**, Orthologous regions of rat, human and mouse genomes encoding pheromone-carrier proteins of the lipocalin family (α_{2u} -globulins in rat and major urinary proteins in mouse) shown in brown. Zfp37-like zinc finger genes are shown in blue. Filled arrows represent likely genes, whereas striped arrows represent likely pseudogenes. Gene expansions are bracketed. Arrowhead orientation represents transcriptional direction. Flanking genes 1 and 2 are *TSCOT* and *CTR1*, respectively. **b**, Site-specific K_A/K_S analysis of rat α_{2u} -globulins. Shown in red are side-chains from codons subject to positive selection. These have been mapped to a ribbon representation of the crystal structure of rat α_{2u} -globulin chain A.

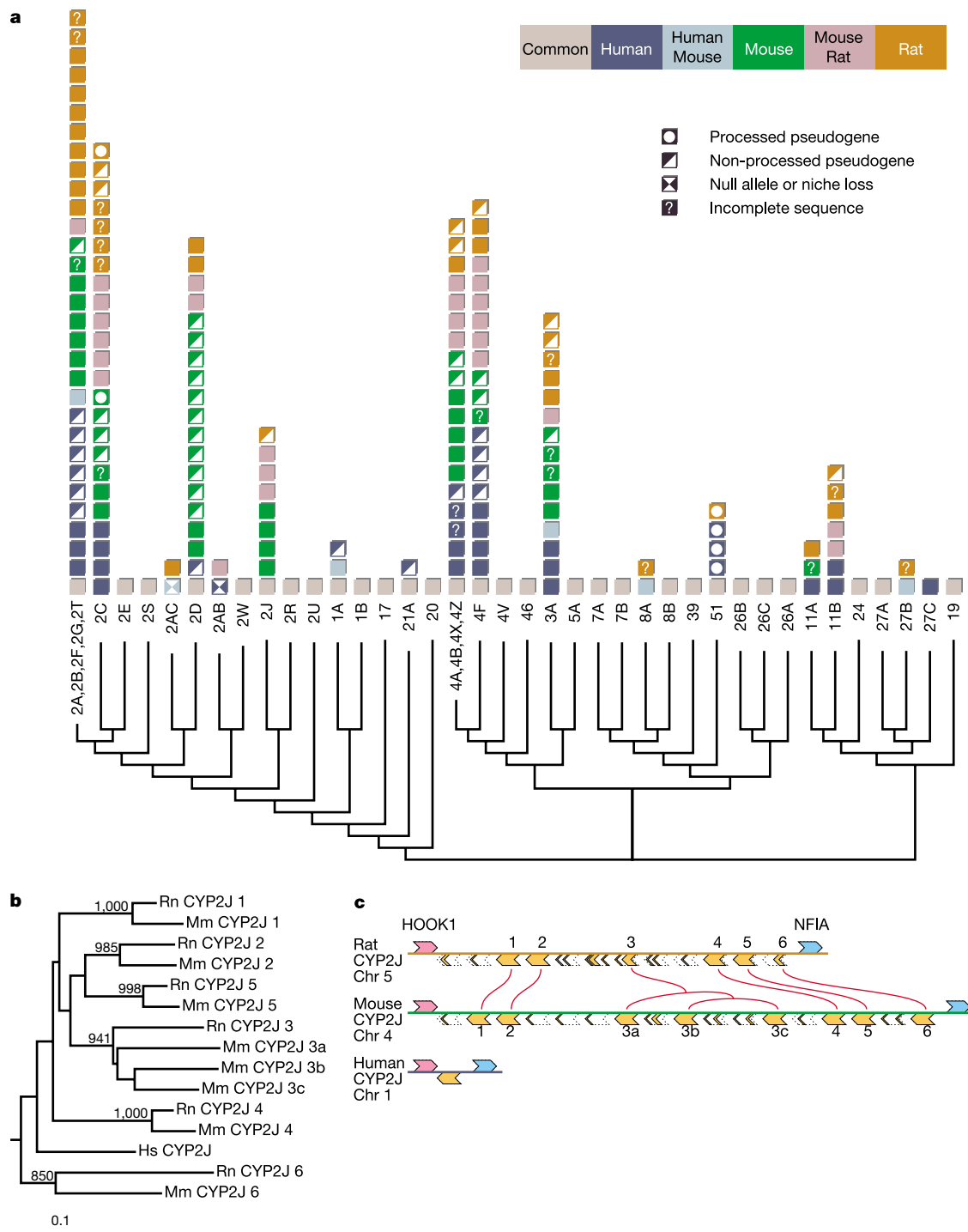


Figure 14 Evolution of cytochrome P450 (CYP) protein families in rat, mouse and human. **a**, Dendrogram topology from 234 full-length sequences. 279 sequences of ≥ 300 amino acids; subfamily names and chromosome numbers are shown. Black branches have $>70\%$ bootstrap support. Incomplete sequences (they contain Ns) are included in counts of functional genes (84 rat, 87 mouse and 57 human) and pseudogenes (including fragments not shown; 77 rat, 121 mouse and 52 human). 64 rat genes and 12 pseudogenes were in predicted gene sets. Human CYP4F is a null allele owing to an in-frame STOP codon in the genome, although a full-length translation exists (SwissProt P98187). Rat CYP27B, missing in the genome, is 'incomplete' because there is a RefSeq entry (NP_446215). Grouped subfamilies CYP2A, 2B, 2F, 2G, 2T and CYP4A, 4B, 4X, 4Z, occur in gene clusters; thus nine loci contain multiple functional genes in a species. One (CYP1A) has fewer rat genes than human, seven have more rodent than human, and all

nine differ in rodent copy numbers. CYP2AC is a rat-specific subfamily (orthologues are pseudogenes). CYP27C has no rodent counterpart. Rodent-specific expansion, rat CYP2J, is illustrated below. **b**, The neighbour-joining tree²²⁴, with the single human gene, contains clear mouse (Mm) and rat (Rn) orthologous pairs (bootstrap values $>700/1,000$ trials shown). Bar indicates 0.1 substitutions per site. **c**, All rat genes have a single mouse counterpart except for CYP2J 3, which has further expanded in mouse (mouse CYP2J 3a, 3b and 3c) by two consecutive single duplications. The genes flanking the CYP2J orthologous regions (rat chromosome 5, 126.9–127.3 Mb; mouse chromosome 4, 94.0–94.6 Mb; human chromosome 1, 54.7–54.8 Mb) are hook1 (HOOK1; pink) and nuclear factor I/A (NFIA; cyan). Genes (solid) and gene fragments (dashed boxes) are shown above (forward strand) and below (reverse strand) the horizontal line. No orthology relation could be concluded for most of these cases.

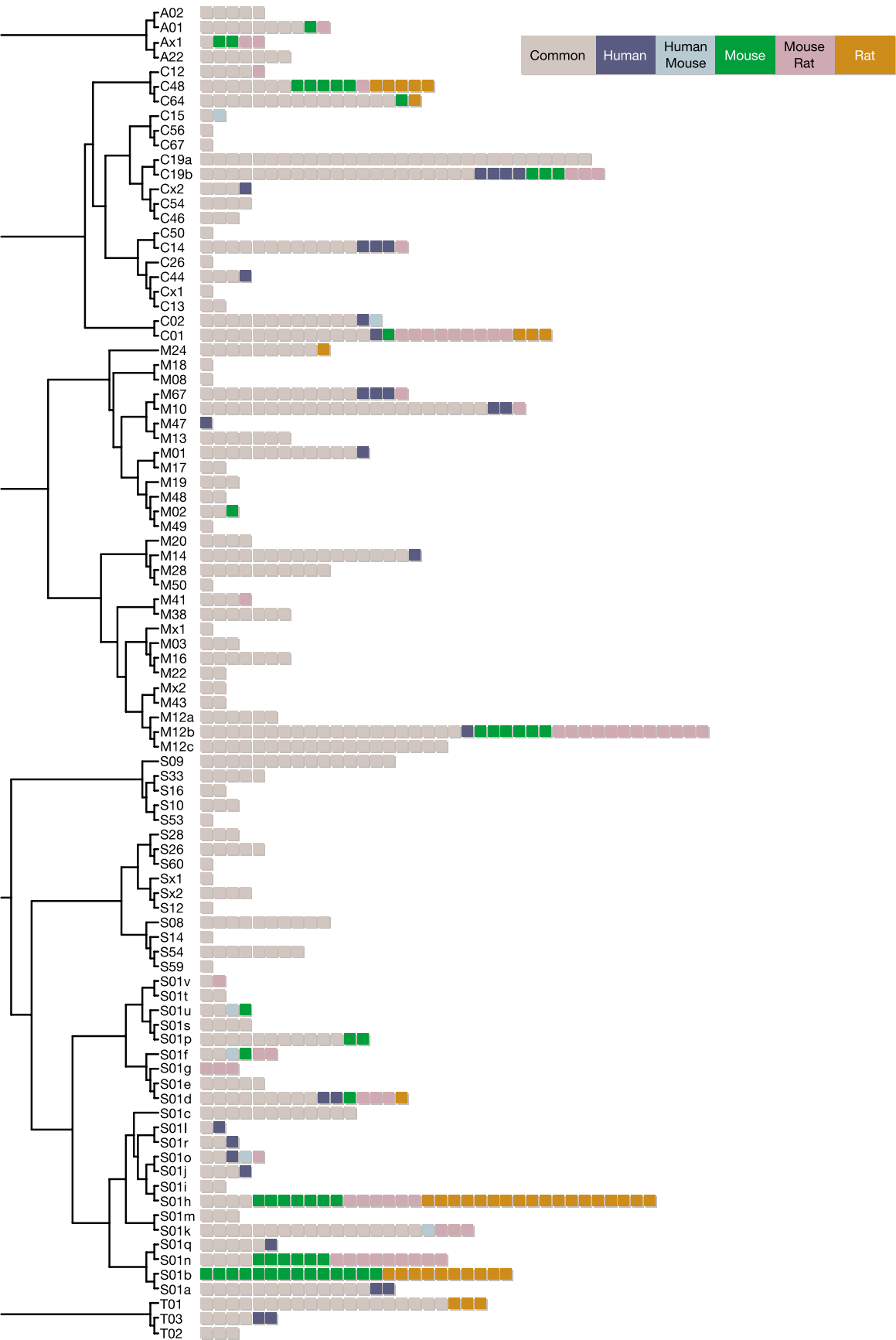


Figure 15 Comparative analysis of rat, mouse and human proteases. The complete non-redundant set of proteases and protease homologues from each species is distributed in five catalytic classes and 67 families. Each square represents a single protease, and is coloured according to its presence or absence in rat, mouse and human as indicated in the inset.

Prevalent, medium-length duplications in rodents

In addition to the transpositionally derived interspersed repeats and simple repeats detected by RepeatMasker and Tandem Repeat Finder, the rat and mouse genomes contain a substantial amount of medium-length unclassified duplications (typically 100–5,000 bp). These are readily seen in self-comparisons and in intra-rodent comparisons after masking the known repeats, but they are substantially less prevalent in comparisons with the human genome (Supplementary Information). Clearly, a substantial fraction of the rodent genomes consists of currently unexplained repeats and a full characterization awaits further studies. The unclassified duplications may include: (1) novel families of low-copy rodent interspersed repeats; (2) extensions of known but not fully characterized rodent repeats; and (3) duplications generated by a mechanism different from transposition.

Rat-specific biology

A principal ambition of the RGSP was to reveal genetic differences between rats and mice that might specify their differences in physiology and behaviour. This view was well supported by the current draft sequence and predicted gene set. In particular, recently duplicated genes are enriched in elements involved in chemosensation and functional aspects of reproduction (Table 5). Here we illustrate the differences in the gene complements of rat and mouse by in-depth analyses of olfactory receptors (ORs), pheromones, cytochromes P450, proteases and protease inhibitors.

Chemosensation

The ability to emit and sense specific smells is a key feature of survival for most animals in the wild. Another paper¹⁸⁰ describes the evolution of rat and mouse pheromones, vomeronasal receptors, and ORs whose genes were duplicated frequently during the time since the common ancestor of rats and mice (Table 5). Their study yielded over 200 aligned codons predicted to have been subject to adaptive evolution. They attribute the rapid evolution of these genes to conspecific competition—in particular, sexual selection.

Using a homology-based identification procedure with manual curation¹⁸¹, we found 1,866 ORs in 113 locations in the rat genome: 69 multi-gene clusters and 44 single genes. After adjusting for missing sequences (the assembly covers 90.2% of the genome), we extrapolate that there are ~2,070 OR genes and pseudogenes. The rat therefore has ~37% more OR genes and pseudogenes than the ~1,510 ORs of the mouse^{181,182}, assuming similar representation of recently duplicated sequences in the two genome assemblies used. Of the 1,774 OR sequences that are not interrupted by assembly gaps, 1,227 (69%) encode intact proteins, while the remaining 547 (31%) sequences are probably pseudogenes with in-frame stop codons, frameshifts, and/or interspersed repeat elements. Fewer mouse OR homologues are pseudogenes (~20%)^{181,182}, but the larger family size in rat still leaves it with substantially more intact ORs than the mouse (~1,430 versus ~1,210). Striking rat-specific expansions of two ancestral clusters account for much of the difference in OR family size and pseudogene content between rat and mouse, although many other clusters exhibit more subtle changes (not shown). Significant differences between human and mouse OR families have also been reported^{181–183}, but the functional implications of OR repertoire size on the ability of different species to detect and discriminate odorants are not yet known.

α_{2u} -globulin pheromones

The α_{2u} -globulin genes are odorant-binding proteins that also contribute to essential survival functions in animals. α_{2u} -globulin homologues are likely to be highly heterogeneous among murid species. Several homologues (major urinary proteins) sequenced from the BALB/c mouse are distinct from their C57BL/6J mouse counterparts, and these also appear to be arranged differently along its genome¹⁸⁴. Moreover, two full-length genes from other mouse

strains¹⁸⁵ differ from their C57BL/6J orthologues—either lacking two of the bases or retaining 20 of the bases that render the C57BL/6J sequences likely to be pseudogenes (not shown).

The evolution of α_{2u} -globulin genes on rat chromosome 5 has clearly driven a significant ‘remodelling’ of this genomic region (Fig. 13a). The orthologous human genomic region contains a single homologue, suggesting that the common ancestor of rodents and human possessed one gene. The genome of C57BL/6J mice contains four homologous genes, and seven pseudogenes, whereas the rat genome contains ten α_{2u} -globulin genes and 12 pseudogenes in a single region (Fig. 13a).

Phylogenetic trees constructed using amino acid, and non-coding DNA, sequences show that, surprisingly, the rat α_{2u} -globulin gene clusters appear to have arisen recently via a rapid burst of gene duplication since the rat–mouse split (Table 5; data not shown). This is consistent with the Rfp37-like zinc-finger-like pseudogene having uniquely ‘hitchhiked’ for virtually all of the rat-specific α_{2u} -globulin gene duplications (Fig. 13a). The sequences of these genes are also evolving rapidly, with median K_A/K_S values of 0.77 and 1.06 for rat and mouse genes, respectively. Amino acid sites that appear to have been subject to adaptive evolution are situated both within the ligand-binding cavity, and on the solvent-exposed periphery of the α_{2u} -globulin structure¹³⁹ (Fig. 13b). This demonstrates how genome analysis can reveal the imprint of adaptive evolution from megabase to single-base levels.

The rapid evolution of these genes, and the remodelling of their genomic regions, can be attributed to the known roles of rat α_{2u} -globulins and mouse major urinary proteins in conspecific competition and sexual selection. These proteins are pheromones and pheromone carriers that are present in large quantities in rodent urine, and act as scent markers indicating dominance and subspecies identity^{186,187}.

Detoxification

Cytochrome P450 is a well-recognized participant in metabolic detoxification, and we also observe rapid evolution within this family. These enzymes metabolize a large number of toxic and endogenous compounds¹⁸⁸ and thus are particularly relevant to clinical and pharmacological studies in humans. As rodents are important model organisms for understanding human drug metabolism, it is important to identify 1:1 orthologues and species-specific expansions and losses¹⁸⁹. Compared with human genes, there are clear expansions of several rodent P450 subfamilies, but there are also significant differences between rat and mouse subfamilies (Fig. 14a). The fastest-evolving subfamily seems to be CYP2J, containing a single gene in human, but at least four in rat and eight in mouse (Fig. 14b, c). CYP2J enzymes catalyse the NADPH-dependent oxidation of arachidonic acid to various eicosanoids, which in turn possess numerous biological activities including modulation of ion transport, control of bronchial and vascular smooth muscle tone, and stimulation of peptide hormone secretion¹⁹⁰. The genomic ordering of genes and their phylogenetic tree indicate an ongoing expansion in the rodents (Fig. 14b, c). This suggests that adaptive evolution has been involved in diversifying their functions. Moreover, detailed study of the nuclear receptors, a highly conserved family of transcription factors, revealed that PXR and CAR, two nuclear receptors regulating CYP genes involved with detoxification¹⁹¹, have the two highest nucleotide substitution rates in their ligand binding domains, whereas SF-1, the nuclear receptor regulating CYP19 (ref. 192), which has not undergone expansion, is more conserved, like other nuclear receptors¹⁹³.

Proteolysis

Protease and protease inhibitor genes also represent an example of rapid evolution in the rat genome. Proteases are a structurally and functionally heterogeneous group of enzymes involved in multiple biological and pathological processes¹⁹⁴. The rat contains 626

protease genes, ~1.7% of the rat gene count¹²⁴, more than human (561) but similar to mouse (641)¹²⁵. Of the rat protease genes, 102 are absent from human, and 42 are absent from mouse (Fig. 15). Several rat gene families have expanded, including placental cathepsins, testases, kallikreins and haematopoietic serine proteases; others appear to have formed pseudogenes in humans (Table 8). These protease families are mainly involved in reproductive or immunological functions, and have evolved independently in the rat and mouse lineages.

The rat protease inhibitor complement contains 183 members, similar to mouse (199) but larger than human (156). As with the protease genes, the rapid evolution in protease inhibitors derives from differential expansions of specific families such as serpins and cystatins. The concomitant expansions in rat and mouse proteases and their inhibitors appear to reflect homeostasis of protein turnover.

These gene family expansions dramatically illustrate how large-scale genomic changes have accompanied species-specific innovation. Positive selection of duplicated genes has afforded the rat an enhanced repertoire of precisely those genes that allow reproductive success despite severe competition from both within its own, and with other, species. This serves as a general illustration of the importance of chemosensation, detoxification and proteolysis in innovation and adaptation.

Human disease gene orthologues in the rat genome

A further strong motivation for sequencing the rat genome was to enhance its utility in biomedical research. Although the rat is already recognized as the premier model for studying the physiological aspects of many human diseases, it has not had as prominent a role in the study of simple genetic disease traits. As more than 1,000 human mendelian disorders now have associated loci and alleles, there is now a tremendous opportunity to link the new knowledge of the rat genome with data from the human disease examples. The precise identification of the rat orthologues of human genes that are mutated in disease creates further opportunities to discover and develop rat models.

Predicted rat genes were compared with 1,112 well-characterized human disease genes¹⁹⁵ that were verified and classified on the basis of pathophysiology (H.H., E.E.W., H.W., K.G.W., H.X., L.G., P.D.S., D.N.C., D.S., M.M.A., C.P.P. and K.F., unpublished work). As predicted by Ensembl, 844 (76%) have 1:1 orthologues in the rat. These predictions are likely to be of high quality because 97.4% of

the 11,422 rat:human 1:1 orthologues predicted by Ensembl were found in orthologous genomic regions.

We asked if these 'disease orthologue' pairs were distinguishable from other rat-human orthologues. Ensembl automatically predicts that 11,522 human genes have rat 1:1 orthologues (corresponding to 46% of all Ensembl predicted human genes). By contrast, a much higher proportion (76%) of human disease genes have Ensembl-predicted rat 1:1 orthologues. Careful analysis of the remaining 268 human genes that were not predicted by Ensembl to show 1:1 orthology indicated that only six of the human disease genes lack likely rat orthologues among genome, cDNA, EST and protein sequences¹⁹⁶. Thus, it appears that, in general, genes involved in human disease are unlikely to have diverged, or to have become duplicated, deleted or lost as pseudogenes, between rat and human (conservation of orthologues discussed above).

We next compared K_S , K_A and the K_A/K_S ratio values of 'disease orthologues' with those of all remaining orthologue pairs. Only the K_S distributions differed significantly¹⁹⁶, suggesting that coding regions of human disease genes and their rat counterparts have mutated more rapidly than the non-disease genes. This might result from factors influencing the specific loci, or the disease genes may characteristically reside in genomic regions that exhibit higher mutation rates.

The disease gene set was next grouped into 16 disease-system categories and analysed using a non-parametric test for K_A/K_S (human/rat)¹⁹⁶ (Fig. 16). Only five disease systems exhibited significant K_A/K_S differences with respect to the remaining samples ($P < 0.05$). Neurological and malformation-syndrome disease categories manifested the lowest median K_A/K_S ratios that are consistent with purifying selection acting on these gene sets. With a comparison of the mean to the mean and standard deviation of the null hypothesis, $[(\text{Mean} - \text{Mean}_0)/\text{Std}_0]$ of -4.63 ($P < 0.0001$), the neurological disease gene set revealed the most evidence for purifying selection of the disease gene categories examined. In contrast, the pulmonary, haematological and immune categories manifested the highest median K_A/K_S ratios, and the genes of the immune system disease category, with a value for $(\text{Mean} - \text{Mean}_0)/\text{Std}_0$ of 4.98 ($P < 0.0001$), show the highest K_A/K_S ratios. These results are consistent with a role for more positive selection, or reduced selective constraints, among these genes.

Where possible, we further considered conservation of these pathophysiology-based gene sets among orthologues of more diverse phyla, including mouse, fish, fly, nematode worm and

Table 8 Protease-expanded gene families and pseudogenes in rat, mouse and human genomes

Protease	Rat gene / locus	Human gene / locus	Mouse gene / locus	Function
Absent genes in assembly	13 from 626 (2.07%)	5 from 561 (0.89%)	5 from 641 (0.78%)	
Expanded families				
Placental cathepsins	10 genes / 17p14	Absent	8 genes / 13B3	Reproduction
Testins	3 genes / 17p14	Absent	3 genes / 13B3	Reproduction
Glandular kallikreins	10 genes / 1q21	Absent	15 genes / 7B2	Reproduction
Mast cell chymases/granzymes	28 genes / 15p13	4 genes / 14q11	17 genes / 14C1	Host defence
Human pseudogenes				
Chymosin	1 gene / 2q34	1 ps / 1p13	1 gene / 3F3	Digestion
Distal intestinal serine proteases	2 genes / 10q12	1 ps / 16p12	2 genes / 17A3	Digestion
Pancreatic elastase	1 gene / 7q35	1 ps / 12q13	1 gene / 15F3	Digestion
Fertilins and reproductive ADAMs	7 genes / various loci	6 ps / various loci	8 genes / various loci	Reproduction
Testases	4 genes / 16q12	3 ps / 8p22	9 genes / 8B1	Reproduction
Testis serine proteases	5 genes / various loci	5 ps / various loci	6 genes / various loci	Reproduction
Implantation serine proteases	2 genes / 10q12	1 ps / 16p13	2 genes / 17A3	Reproduction
Airway trypsin-like proteases	3 genes / 14p21	3 ps / 4q13	3 genes / 5E1	Host defence
Rat pseudogenes				
Calpain 13	1 ps / 6q12	1 gene / 2p23	1 gene / 17E2	Reproduction ?
Pyroglutamyl-peptidase II	1 ps / 1q22	1 gene / 15q26	1 gene / 7C	Metabolism
Gln-fructose-6-P transamidase 3	1 ps / Xq21	1 gene / Xq21	1 ps / XC3	Metabolism
Aminopeptidase MAMS/L-RAP	1 ps / 1q12	1 gene / 5q15	1 ps / 17A3	Host defence
Carboxypeptidase O	1 ps / 9q31	1 gene / 2q33	1 ps / 1C2	Unknown
Procollagen III N-endopeptidase	1 ps / 19q12	1 gene / 16q24	1 ps / 8E2	Metabolism ?
Kallikrein-2 and -3	2 ps / 1q21	2 genes / 19q13	1 ps / 7B2	Reproduction
Testis-specific protein 50	1 ps / 8q32	1 gene / 3p21	1 gene / 9F2	Reproduction

yeast orthologues. Overall, we obtained results consistent with those reported here for these rat:human 1:1 orthologous gene disease categories¹⁹⁶. These results demonstrate that the individual genes that constitute various disease systems exhibit significantly different average evolutionary rates. The higher evolutionary rates noted for the immune system disease genes are consistent with a previous finding that lymphocyte-specific genes evolve relatively rapidly¹⁹⁷ and may indicate rapid diversification of the functions of the immune systems of rodents and humans. This is expected for genes involved in controlling species-restricted infectious agents if strong adaptive pressure acts during host–pathogen co-evolution. Thus, the results of studies of these rodent genes may be less directly relevant to our understanding of human immune system diseases than results obtained for other pathophysiology disease systems where conservation is greater and purifying selection is stronger.

We have also specifically examined a number of genes that harbour triplet nucleotide repeats, and are involved in human neurological disorders such as Huntington's disease, a condition known to be caused by CAG triplet repeat expansion producing abnormally long polyglutamine tracts in an otherwise normal protein¹⁹⁸. Analysis of the rat–human orthologues of these disease genes indicated that repeat-expansion disease genes exhibit a repeat length that is substantially shorter in the rat than that found in the normal human gene (Fig. 17). In all cases, human disease genes localize below the line demarcating 1:1 length correlation, showing that rat orthologues uniformly bear shorter repeats. At present, there are no naturally occurring rat strains described that exhibit neurological disease associated with repeat-expansion mechanisms. The shorter repeat length of these orthologues in the rat would be consistent with either the lack of repeat-expansion mutational mechanisms in the rat or the failure of these orthologues to achieve a 'critical repeat length' susceptible to such mutational mechanisms. Other human genes, not at present known to be associated with disease, also contain glutamine repeats that are much shorter in the rat orthologues, and thus, could be investigated as potential disease candidates¹⁹⁶. These triplet-repeat-bearing genes may be susceptible to mutations that arise through repeat-expansion mechanisms. In Fig. 17, it may also be observed that a relatively high proportion of repeats are significantly longer in the rat than in their corresponding human orthologue.

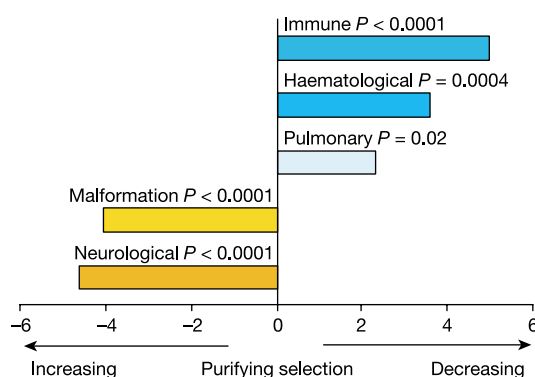


Figure 16 Selective constraints differ for human disease systems in the rat genome. Human disease system categories showing significant differences ($P < 0.05$) in a non-parametric test (Mann–Whitney–Wilcoxon) comparing K_A/K_S (human:rat) ratios. P values from two-level tests between genes from one disease system and the remaining genes. (Mean–Mean0)/Std0 values from multi-level tests from 16 categorized disease systems. Negative values (shown in yellow and orange) for neurological (–4.63) and malformation-syndrome (–4.04) categories were observed to be consistent with K_A/K_S ranges in which purifying selection predominates. Immune, haematological and pulmonary categories show positive values of 4.98, 3.59 and 2.34, respectively (for complete data set and details, see ref. 199).

In addition to enabling the direct comparison of rat–human disease orthologues, the rat genome sequence itself is an invaluable aid for the discovery of additional rat genes that can be studied as disease models. Two general modes can now be pursued. First, genes underlying disease phenotypes with simple inheritance that have been mapped to chromosomal regions can be more easily pursued in both species. Indeed, the rearrangements of conserved segments between the two species in this map were found to have significant value, because they tighten the boundaries of the mapped disease regions and thus reduce the number of genes that could potentially be associated with a given disease phenotype¹¹³. Second, the identification of multiple alleles contributing to quantitative and complex trait differences that are involved in disease processes can be pursued with more accuracy, both in the initial association phases, and in subsequent efforts to detect causative alleles.

Rat single nucleotide polymorphisms

The discovery and cataloguing of the natural DNA variation that persists between individual rat strains will allow further research using rat model systems. Although many rat microsatellites have been characterized and studied, single nucleotide polymorphisms (SNPs) are of more general interest because of their probable ubiquity, and the ease with which they can be assayed. SNP data have three broad applications: (1) the individual markers can be used in ongoing efforts to associate phenotypes that have complex underlying genetic components, with specific sites in the genome. (2) A panel of such markers can be used in conjunction with selective breeding and chromosome mechanics, to generate rat strains that are amenable to the kinds of manipulations that will hasten the discovery of important alleles. (3) A set of such markers can be used to detail the history of the different genomic events that have led to the structure of the genomes of contemporary rat strains. A detailed map of these events has a utility analogous to the current human haplotype (HapMap) mapping project¹⁹⁹ and will probably

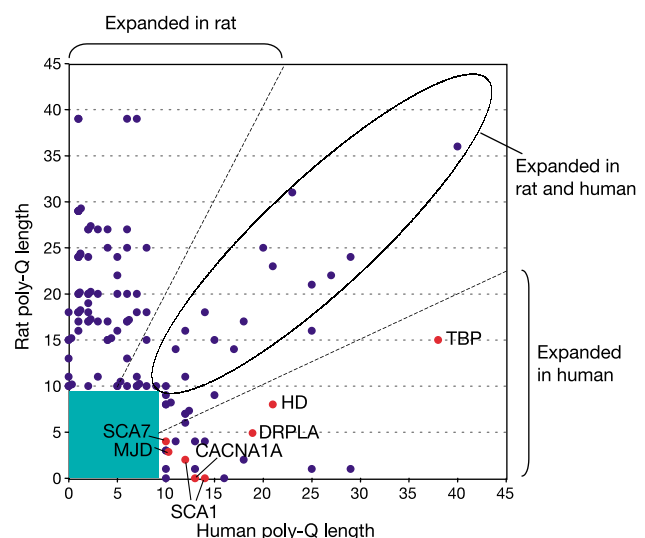


Figure 17 Polyglutamine repeat length comparison between human and rat. Points represent protein poly-Q length for rat and human. Red points correspond to repeats in genes associated with human disease: SCA1, spinocerebellar ataxia 1 protein, or ataxin1; SCA7, spinocerebellar ataxia 7 protein; MJD, Machado–Joseph disease protein; CACNA1A, spinocerebellar ataxia 6 protein, or calcium channel alpha 1A subunit isoform 1; DRPLA, dentatorubral pallidoluysian atrophy protein; HD, Huntington's disease protein, or huntingtin; TBP, TATA binding protein or spinocerebellar ataxia 17 protein. Repeat lengths over ten were examined; green shading delineates the range not included in our analysis. Also noted are a set that are expanded in rat and human (black circle) and a set where repeats are expanded in the rat.

aid disease gene identification, as recently suggested for the mouse²⁰⁰.

The Rnor3.1 draft sequence was generated primarily from DNA of a single inbred rat line. This maximized the likelihood of deriving an accurate sequence assembly, but reduced any likely discovery of natural variation in this phase of the project. As a consequence there has been no large-scale public SNP discovery from rat genomic sequencing. A pilot project based on coding (c)SNP discovery has been initiated, however²⁰¹, as these cSNPs represent a particularly important subset of variants that may have direct functional significance²⁰². These data have illustrated both immediate applications and the long-term potential for an effort aimed at comprehensive SNP discovery.

Conclusions

As the third mammalian genome to be sequenced, the rat genome has provided both predictable and surprising information about mammalian species. Although it was clear at the outset of this programme that ongoing rat research would benefit from the resource of a genome sequence, there was uncertainty about how many new insights would be found, especially considering the superficial similarities between the rat and the already sequenced mouse. Instead, the results of the sequencing and analysis have generated some deep insights into the evolutionary processes that have given rise to these different species. In addition, the project has been invaluable in further developing the methods for the generation and analysis of large genome sequence data sets.

The generation of the rat draft tested the new 'combined approach' for large genome sequencing. As the overall assembly is of high quality, there is no doubt that this overall strategy, and the supporting software we have developed, provides a suitable approach for this problem. Because we included a BAC 'skimming' component in the underlying data set, the assembly recovered a fraction of the genome that was expected, by analogy to the mouse project, to be difficult to assemble from pure WGS data. In addition, the BAC skimming component allowed progressive generation of high-quality local assemblies that were of use to the rat research community as the project developed. On the other hand, although the BAC component used here was far less expensive than the fully ordered and highly redundant set used in the hierarchical approach to sequencing the human genome, it nevertheless increased the overall cost of data production relative to a WGS approach.

The issue of efficacy of WGS versus other approaches to the sequencing of large genomes remains a matter of earnest scientific debate. In ongoing projects at different centres that participated in the RGSP consortium, different approaches are being used to tackle new genomes. These include pure WGS methods, the combined approach and variations on that methodology. The future application of the different procedures depends on the target genome sizes, the expected degree of heterogeneity (that is, polymorphism) in the organism to be sequenced, and the preferences of the individual centre. So far, all the genomes that have been analysed by RGSP consortium members have been of high quality and we anticipate that this will continue as the benefits and disadvantages of different approaches are further studied and analysed.

The rat genome data have improved the utility of the rat model enormously. Now that near-complete knowledge of the rat gene content is realizable, individual researchers have a data source for the rat 'parts list' that can be explored with the high degree of confidence and precision that is appropriate for biomedical research. A similar improvement has been made in the resources for physical and genetic mapping, because the relative position of individual markers is now known with high confidence and there are now computational resources to bridge the process of genetic association with gene modelling and experimental investigation. These advances have been reflected by measured increases in the use of all the rat-specific public genome data sets that can be accessed

online, as well as by the informally assessed increases in overall 'genomic' research of this model.

The expected benefit of a third mammalian sequence providing an outgroup by which to discriminate the timing of events that had already been noted between mouse and human was fully realized. Using the three sequences and other partial data sets from additional organisms, it was possible to measure some of the overall faster rate of evolutionary change in the rodent lineage shared by mice and rats, as well as the peculiar acceleration of some aspects of rat-specific evolution. The observation of specific expanded gene families in the rat should provide material for targeted studies for some time.

At this time there is no plan to further upgrade or finish the rat genome sequence. This programme decision is a consequence of the high cost of converting draft sequence to finished data, and the pressing need to analyse new genomes. However, as the distant objective of very-low-cost sequencing or other advances that can improve draft sequences inexpensively are realized, it might be envisioned that a rat sequence that approaches the quality of the current human data will be produced. A finished rat genome may answer many questions, as specific clues already show that areas of the genome that are most difficult to resolve in a random sequencing project are also those areas that are most dynamic, and therefore of high potential interest in an evolutionary context.

Despite the advances represented here, we are clearly still at the beginning of the full analysis of the mammalian genome and its complex evolutionary history. Much of the additional data that are required to complete this story will be from other genomes, distantly related to rat. Nevertheless, a considerable body of data remains to be developed from this species. In addition to the distant prospect of a finished rat genome, analysis of other rat strains may yield genome-wide polymorphism data, while targeted efforts to generate cDNA clone collections will provide rat-specific reagents for routine use in research. Together with the ongoing efforts to fully develop methods to genetically manipulate whole rats and provide effective 'gene knockouts', the current and future rat genome resources will ensure a place for this organism in genomic and biomedical research for some time. □

Methods

DNA sequencing and data access

Paired-end reads from BAC and WGS libraries were produced as previously described^{2,203}. Unprocessed sequence reads are available from the NCBI Trace Archive (ftp://ftp.ncbi.nih.gov/pub/TraceDB/rattus_norvegicus/); raw eBAC assembly data are available from the BCM-HGSC (<http://www.hgsc.bcm.tmc.edu/Rat/>); and the released Rnor3.1 assembly is available from the BCM-HGSC (<ftp://ftp.hgsc.bcm.tmc.edu/pub/analysis/rat/>), the NCBI (ftp://ftp.ncbi.nih.gov/genomes/R_norvegicus/), and the UCSC (<http://genome.ucsc.edu/downloads.html>).

Genome assembly

Assembly of the rat genome by the *Atlas* system is described in detail elsewhere⁵⁴. Earlier assemblies (Rnor2.0/2.1) of the initial data set were based on 40 million total reads and 19,000 BAC skims. These assemblies spanned 2.66 Gb and comprised over 900 ultrabactigs with N_{50} of over 5 Mb. They differed only in the removal of short artefactual duplications from Rnor2.0. Rnor3.1 includes another 1,100 BACs, selected to fill gaps in Rnor2.1. Because of the comprehensive coverage of the genome by Rnor2.0/2.1, it was used for the initial predictions of genes and proteins.

BAC fingerprints

An agarose-gel-based fingerprinting methodology^{204–207} was employed to generate *Hind*III fingerprints from 199,782 clones in the CHORI-230 BAC library. The contig assembly was subjected to manual review and editing to refine clone order within contigs and to make merges between contigs, using tools provided in the FPC software^{208–210}. Fingerprints for 5,250 RPCI-31 PACs²¹¹ and RPCI-32 BACs were subsequently added to allow correlation between the fingerprint map and a developing YAC map of the rat genome. BAC and PAC clones are available through BACPAC Resources at CHORI (bacpacorders@chori.org).

BAC, PAC and YAC maps

Markers generated from BAC and PAC clones were hybridized against YAC⁵⁸ (R.D., Pmatch, unpublished software) and radiation hybrid libraries^{61,212} to produce independent maps that were subsequently combined. Genetic markers from two rat

genetic maps⁶¹ and the radiation hybrid map⁵⁹ were aligned to the Rnor3.1 assembly using BLAT¹²³ (when sequence was available) or electronic polymerase chain reaction (EPCR)²¹³.

Finished sequence used for quality assessment of the assembly

To assess the accuracy of the *Atlas* assembly, the Rnor3.1 sequence was compared to 13 Mb of sequences that had been finished to high quality.

Large-scale rearrangements

We compared these assemblies: Human (April 2003, NCBI build 33); Mouse (February 2003, NCBI build 30); and Rat (June 2003, Rnor3.1). Repeats were masked using RepeatMasker (A.S. & P. Green, unpublished work; see <http://ftp.genome.washington.edu/RM/RepeatMasker.html>) and TandemRepeatFinder²¹⁴. Local alignments were produced using PatternHunter⁷⁰ (Supplementary Information). Repeat contamination was removed and the remaining similarities combined into two- and three-way anchors⁷³ and synteny blocks produced at various resolutions using GRIMM-Synteny⁷¹.

Genome-wide visualization of conserved synteny

Pairwise comparisons of the genomes of human, mouse and rat using MULTIZ^{69,215}, MLAGAN^{216,217}, MAVID¹¹⁰, PatternHunter⁷⁰ and Pash⁷² were merged into blocks of conserved synteny^{69,71,72}, and the 1-Mb-resolution images were displayed using the Virtual Genome Painting method (M.L.G.-G. *et al.*, unpublished work; <http://www.genbore.org>).

Rat segmental duplications

Segmental duplications >5 kb were identified, extracted and aligned as described²¹⁸, and paralogous sequence relationships were assessed using PARASIGHT visualization software (J.A.B., unpublished work; Supplementary Information).

Venn diagram

Pairwise and three-way alignments generated using BLASTZ²¹⁹ and MULTIZ²¹⁵ or HUMOR²¹⁵ were analysed to classify each nucleotide in the three genomes by the species with which it aligns: in all three species, aligning between human and rat (but not mouse), between human and mouse (but not rat), or between mouse and rat (but not human). Other nucleotides are species-specific; unassigned nucleotides occupying gaps in the genome assemblies were excluded. On the basis of output from RepeatMasker¹⁶⁴ and RepeatDate⁸⁹, nucleotides were assigned to categories (of non-repetitive, repetitive with a certain ancestry, or repetitive but unassigned) and counted. See Supplementary Table SI-1 for details.

Gene prediction

ENSEMBL transcript models were built from 28,478 rodent proteins that were aligned to the genome using a combination of Pmatch (R.D., unpublished software), BLAST²²⁰ and GeneWise²²¹. Models based on 5,083 vertebrate proteins were added in regions without rodent-protein-based models. UTRs were added using 11,170 transcripts built from 8,615 different rat cDNAs aligned to the genome using BLAT, with coverage $\geq 90\%$ and identity $\geq 95\%$. This procedure (as described¹¹² but without GENSCAN predictions), gave rise to 18,241 genes and 20,373 transcripts. This is the protein-based gene set. Rat and mouse cDNA and rat EST-based gene sets were also built. See Supplementary Information for details.

Non-processed pseudogene identification

Human and mouse genes related by 1:1 orthology and lacking an apparent rat orthologue were considered. See Supplementary Information for details.

High-resolution analyses of chromosome 10

These were performed predominantly on the whole genome alignments²¹⁷. Plots in Fig. 9 were generated by sliding windows of width 2 Mb and a step size of 400 kb (total = 277 windows). See Supplementary Information for details.

Received 31 December 2003; accepted 20 February 2004; doi:10.1038/nature02426.

1. Darwin, C. On *The Origin of Species by Means of Natural Selection* 1st edn, Ch. 4, 108 (John Murray, London, 1859).
2. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
3. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
4. Adkins, R. M., Gelke, E. L., Rowe, D. & Honeycutt, R. L. Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol. Biol. Evol.* **18**, 777–791 (2001).
5. Springer, M. S., Murphy, W. J., Eizirik, E. & O'Brien, S. J. Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc. Natl Acad. Sci. USA* **100**, 1056–1061 (2003).
6. Canby, T. Y. The rat, lapdog of the devil. *Nat. Geogr.* July, 60–87 (1977).
7. Robinson, R. *Genetics of the Norway Rat* (Pergamon, Oxford, 1965).
8. Barnett, S. A. *The Story of Rats. Their Impact on Us, and Our Impact on Them* Ch. 2, 17–18 (Allen and Unwin, Crows Nest, Australia, 2002).
9. Hedrich, H. J. in *History, Strains, and Models in the Laboratory Rat* (ed. Krinke, G. J.) 3–16 (Academic, San Diego, 2000).
10. Lindsey, J. R. in *The Laboratory Rat* (eds Baker, H. J., Lindsey, J. R. & Weisbroth, S. H.) 1–36 (Academic, New York, 1979).
11. Greenhouse, D. D., Festing, M. F. W., Hasan, S. & Cohen, A. L. in *Genetic Monitoring of Inbred Strains of Rats* (ed. Hedrich, H. J.) 410–480 (Gustav Fischer, Stuttgart, 1990).

12. Kuntz, C. *et al.* Comparison of laparoscopic versus conventional technique in colonic and liver resection in a tumor-bearing small animal model. *Surg. Endosc.* **16**, 1175–1181 (2002).
13. Kitagawa, K., Hamada, Y., Nakai, K., Kato, Y. & Okumura, T. Comparison of one- and two-step procedures in a rat model of small bowel transplantation. *Transplant. Proc.* **34**, 1030–1032 (2002).
14. Sauve, Y., Girmar, S. V., Wang, S., Keegan, D. J. & Lund, R. D. Preservation of visual responsiveness in the superior colliculus of RCS rats after retinal pigment epithelium cell transplantation. *Neuroscience* **114**, 389–401 (2002).
15. Wang, H. *et al.* Attenuation of acute xenograft rejection by short-term treatment with LF15–0195 and monoclonal antibody against CD45RB in a rat-to-mouse cardiac transplantation model. *Transplantation* **75**, 1475–1481 (2003).
16. Alves, A. *et al.* Total vascular exclusion of the liver enhances the efficacy of retroviral-mediated associated thymidine kinase and interleukin-2 genes transfer against multiple hepatic tumors in rats. *Surgery* **133**, 669–677 (2003).
17. Liu, M. Y., Poellinger, L. & Walker, C. L. Up-regulation of hypoxia-inducible factor 2 α in renal cell carcinoma associated with loss of Tsc-2 tumor suppressor gene. *Cancer Res.* **63**, 2675–2680 (2003).
18. Jin, X. *et al.* Effects of leptin on endothelial function with OB-Rb gene transfer in Zucker fatty rats. *Atherosclerosis* **169**, 225–233 (2003).
19. Ravingerova, T., Neckar, J. & Kolar, F. Ischemic tolerance of rat hearts in acute and chronic phases of experimental diabetes. *Mol. Cell. Biochem.* **249**, 167–174 (2003).
20. Taylor, J. R. *et al.* An animal model of Tourette's syndrome. *Am. J. Psychiatry* **159**, 657–660 (2002).
21. Smyth, M. D., Barbaro, N. M. & Baraban, S. C. Effects of antiepileptic drugs on induced epileptiform activity in a rat model of dysplasia. *Epilepsy Res.* **50**, 251–264 (2002).
22. McBride, W. J. & Li, T. K. Animal models of alcoholism: neurobiology of high alcohol-drinking behavior in rodents. *Crit. Rev. Neurobiol.* **12**, 339–369 (1998).
23. Crisci, A. R. & Ferreira, A. L. Low-intensity pulsed ultrasound accelerates the regeneration of the sciatic nerve after neurotomy in rats. *Ultrasound Med. Biol.* **28**, 1335–1341 (2002).
24. Ozkan, O. *et al.* Reinnervation of denervated muscle in a split-nerve transfer model. *Ann. Plast. Surg.* **49**, 532–540 (2002).
25. Fray, M. J., Dickinson, R. P., Huggins, J. P. & Occlleston, N. L. A potent, selective inhibitor of matrix metalloproteinase-3 for the topical treatment of chronic dermal ulcers. *J. Med. Chem.* **46**, 3514–3525 (2003).
26. Petratos, P. B. *et al.* Full-thickness human foreskin transplantation onto nude rats as an *in vivo* model of acute human wound healing. *Plast. Reconstr. Surg.* **111**, 1988–1997 (2003).
27. Hussar, P. *et al.* Bone healing models in rat tibia after different injuries. *Ann. Chir. Gynaecol.* **90**, 271–279 (2001).
28. Yang, T. D., Pei, J. S., Yang, S. L., Liu, Z. Q. & Sun, R. L. Medical prevention of space motion sickness—animal model of therapeutic effect of a new medicine on motion sickness. *Adv. Space Res.* **30**, 751–755 (2002).
29. Forte, A. *et al.* Stenosis progression after surgical injury in Milan hypertensive rat carotid arteries. *Cardiovasc. Res.* **60**, 654–663 (2003).
30. Komamura, K. *et al.* Differential gene expression in the rat skeletal and heart muscle in glucocorticoid-induced myopathy: analysis by microarray. *Cardiovasc. Drugs Ther.* **17**, 303–310 (2003).
31. McBride, M. W. *et al.* Functional genomics in rodent models of hypertension. *J. Physiol. (Lond.)* **554**, 56–63 (2004).
32. Kasteleijn-Nolst Trenite, D. G. & Hirsch, E. Levetiracetam: preliminary efficacy in generalized seizures. *Epileptic Disord.* **5**, S39–S44 (2003).
33. Malik, A. S. *et al.* A novel dehydroepiandrosterone analog improves functional recovery in a rat traumatic brain injury model. *J. Neurotrauma* **20**, 463–476 (2003).
34. Kostubsky, V. E. *et al.* Evaluation of hepatotoxic potential of drugs by inhibition of bile acid transport in cultured primary human hepatocytes and intact rats. *Toxicol. Sci.* **76**, 220–228 (2003).
35. Lindon, J. C. *et al.* Contemporary issues in toxicology: the role of metabolomics in toxicology and its evaluation by the COMET project. *Toxicol. Appl. Pharmacol.* **187**, 137–146 (2003).
36. Tam, R. C. *et al.* The ribavirin analog ICN 17261 demonstrates reduced toxicity and antiviral effects with retention of both immunomodulatory activity and reduction of hepatitis-induced serum alanine aminotransferase levels. *Antimicrob. Agents Chemother.* **44**, 1276–1283 (2000).
37. Youssef, A. F., Turck, P. & Fort, F. L. Safety and pharmacokinetics of oral lansoprazole in preadolescent rats exposed from weaning through sexual maturity. *Reprod. Toxicol.* **17**, 109–116 (2003).
38. National Institutes of Health. *Network for Large-Scale Sequencing of the Rat Genome* (<http://grants2.nih.gov/grants/guide/rfa-files/RFA-HG-00-002.html>) (2000).
39. Aparicio, S. *et al.* Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301–1310 (2002).
40. Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92 (2002).
41. Goff, S. A. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100 (2002).
42. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
43. Dehal, P. *et al.* The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* **298**, 2157–2167 (2002).
44. Myers, E. W. *et al.* A whole-genome assembly of *Drosophila*. *Science* **287**, 2196–2204 (2000).
45. Myers, E. W., Sutton, G. G., Smith, H. O., Adams, M. D. & Venter, J. C. On the sequencing and assembly of the human genome. *Proc. Natl Acad. Sci. USA* **99**, 4145–4146 (2002).
46. Waterston, R. H., Lander, E. S. & Sulston, J. E. On the sequencing of the human genome. *Proc. Natl Acad. Sci. USA* **99**, 3712–3716 (2002).
47. Waterston, R. H., Lander, E. S. & Sulston, J. E. More on the sequencing of the human genome. *Proc. Natl Acad. Sci. USA* **100**, 3022–3024 (2003); author reply (100), 3025–3026 (2003).
48. Green, P. Whole-genome disassembly. *Proc. Natl Acad. Sci. USA* **99**, 4143–4144 (2002).
49. Batzoglou, S. *et al.* ARACHNE: a whole-genome shotgun assembler. *Genome Res.* **12**, 177–189 (2002).

50. Jaffe, D. B. *et al.* Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res.* **13**, 91–96 (2003).
51. Cheung, J. *et al.* Recent segmental and gene duplications in the mouse genome. *Genome Biol.* **4**, R47 [online] (2003).
52. Eichler, E. E. Masquerading repeats: Paralogous pitfalls of the human genome. *Genome Res.* **8**, 758–762 (1998).
53. Eichler, E. E. Segmental duplications: what's missing, misassigned, and misassembled—and should we care? *Genome Res.* **11**, 653–656 (2001).
54. Havlak, P. *et al.* The Atlas genome assembly system. *Genome Res.* **14**, 721–732 (2004).
55. Osogawa, K. *et al.* BAC Resources for the rat genome project. *Genome Res.* **14**, 780–785 (2004).
56. Krzywinski, M. *et al.* Integrated and sequence-ordered BAC and YAC-based physical maps for the rat genome. *Genome Res.* **14**, 766–779 (2004).
57. Chen, R., Sodergren, E., Gibbs, R. & Weinstock, G. M. Dynamic building of a BAC clone tiling path for genome sequencing project. *Genome Res.* **14**, 679–684 (2004).
58. Cai, L. *et al.* Construction and characterization of a 10-genome equivalent yeast artificial chromosome library for the laboratory rat, *Rattus norvegicus*. *Genomics* **39**, 385–392 (1997).
59. Kwitek, A. E. *et al.* High density rat radiation hybrid maps containing over 24,000 SSLPs, genes, and ESTs provide a direct link to the rat genome sequence. *Genome Res.* **14**, 750–757 (2004).
60. Felsenfeld, A., Peterson, J., Schloss, J. & Guyer, M. Assessing the quality of the DNA sequence from the Human Genome Project. *Genome Res.* **9**, 1–4 (1999).
61. Steen, R. G. *et al.* A high-density integrated genetic linkage and radiation hybrid map of the laboratory rat. *Genome Res.* **9** (insert), AP1–AP8 (1999).
62. Misra, S. *et al.* Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol.* **3**, RESEARCH0083.1–0083.22 [online] (2002).
63. Li, X. & Waterman, M. S. Estimating the repeat structure and length of DNA sequences using L-tuples. *Genome Res.* **13**, 1916–1922 (2003).
64. Riethman, H. *et al.* Mapping and initial analysis of human subtelomeric sequence assemblies. *Genome Res.* **14**, 18–28 (2004).
65. Bayona-Bafaluy, M. P. *et al.* Revisiting the mouse mitochondrial DNA sequence. *Nucleic Acids Res.* **31**, 5349–5355 (2003).
66. Kent, W. J., Baertsch, R., Hinrichs, A., Miller, W. & Haussler, D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proc. Natl Acad. Sci. USA* **100**, 11484–11489 (2003).
67. Pevzner, P. & Tesler, G. Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proc. Natl Acad. Sci. USA* **100**, 7672–7677 (2003).
68. Nadeau, J. H. & Taylor, B. A. Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc. Natl Acad. Sci. USA* **81**, 814–818 (1984).
69. Schwartz, S. *et al.* Human–mouse alignments with BLASTZ. *Genome Res.* **13**, 103–107 (2003).
70. Ma, B., Tromp, J. & Li, M. PatternHunter: faster and more sensitive homology search. *Bioinformatics* **18**, 440–445 (2002).
71. Pevzner, P. & Tesler, G. Genome rearrangements in mammalian evolution: lessons from human and mouse genomes. *Genome Res.* **13**, 37–45 (2003).
72. Kalafus, K. J., Jackson, A. R. & Milosavljevic, A. Pash: Efficient genome-scale sequence anchoring by positional hashing. *Genome Res.* **14**, 672–678 (2004).
73. Bourque, G., Pevzner, P. A. & Tesler, G. Reconstructing the genomic architecture of ancestral mammals: lessons from human, mouse, and rat genomes. *Genome Res.* **14**, 507–516 (2004).
74. Graves, J. A., Gecz, J. & Hameister, H. Evolution of the human X—a smart and sexy chromosome that controls speciation and development. *Cytogenet. Genome Res.* **99**, 141–145 (2002).
75. Bourque, G. & Pevzner, P. A. Genome-scale evolution: reconstructing gene orders in the ancestral species. *Genome Res.* **12**, 26–36 (2002).
76. Murphy, W. J., Bourque, G., Tesler, G., Pevzner, P. & O'Brien, S. J. Reconstructing the genomic architecture of mammalian ancestors using multispecies comparative maps. *Hum. Genom.* **1**, 30–40 (2003).
77. Kirkness, E. F. *et al.* The dog genome: survey sequencing and comparative analysis. *Science* **301**, 1898–1903 (2003).
78. Murphy, W. J., Sun, S., Chen, Z. Q., Pecon-Slattery, J. & O'Brien, S. J. Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. *Genome Res.* **9**, 1223–1230 (1999).
79. Ventura, M., Archidiacono, N. & Rocchi, M. Centromere emergence in evolution. *Genome Res.* **11**, 595–599 (2001).
80. Yunis, J. J. & Prakash, O. The origin of man: a chromosomal pictorial legacy. *Science* **215**, 1525–1530 (1982).
81. Murphy, W. J., Fronicke, L., O'Brien, S. J. & Stanyon, R. The origin of human chromosome 1 and its homologs in placental mammals. *Genome Res.* **13**, 1880–1888 (2003).
82. Stanyon, R., Stone, G., Garcia, M. & Froenicke, L. Reciprocal chromosome painting shows that squirrels, unlike murid rodents, have a highly conserved genome organization. *Genomics* **82**, 245–249 (2003).
83. Bailey, J. A. *et al.* Recent segmental duplications in the human genome. *Science* **297**, 1003–1007 (2002).
84. Thomas, J. W. *et al.* Pericentromeric duplications in the laboratory mouse. *Genome Res.* **13**, 55–63 (2003).
85. Horvath, J. E. *et al.* Using a pericentromeric interspersed repeat to recapitulate the phylogeny and expansion of human centromeric segmental duplications. *Mol. Biol. Evol.* **20**, 1463–1479 (2003).
86. Guy, J. *et al.* Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10p. *Genome Res.* **13**, 159–172 (2003).
87. Tuzun, E., Bailey, J. A. & Eichler, E. E. Recent segmental duplications in the working draft assembly of the brown Norway rat. *Genome Res.* **14**, 493–506 (2004).
88. Pruitt, K. D. & Maglott, D. R. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.* **29**, 137–140 (2001).
89. Yang, S. *et al.* Patterns of insertions and their covariation with substitutions in the rat, mouse and human genomes. *Genome Res.* **14**, 517–527 (2004).
90. Roskin, K. M., Diekhans, M. & Haussler, D. In *Proc. 7th Annu. Int. Conf. Res. Comput. Mol. Biol. (RECOMB 2003)* (eds Vingron, M., Istrail, S., Pevzner, P. & Waterman, M.), doi:10.1145/640075.640109, 257–266 (ACM Press, New York, 2003).
91. Chiaromonte, F. *et al.* The share of human genomic DNA under selection estimated from human–mouse genomic alignments. *Cold Spring Harbor Symp. Quant. Biol.* (in the press).
92. Cooper, G. M. *et al.* Characterization of evolutionary rates and constraints in three mammalian genomes. *Genome Res.* **14**, 539–548 (2004).
93. Dermitzakis, E. T. *et al.* Numerous potentially functional but non-genic conserved sequences on human chromosome 21. *Nature* **420**, 578–582 (2002).
94. Dermitzakis, E. T. *et al.* Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs). *Science* **302**, 1033–1035 (2003).
95. Nekrutenko, A. Rat–mouse comparisons to identify rodent-specific exons. *Genome Res.* (in the press).
96. Hardison, R. C. *et al.* Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. *Genome Res.* **13**, 13–26 (2003).
97. Thomas, J. W. *et al.* Comparative analyses of multi-species sequences from targeted genomic regions. *Nature* **424**, 788–793 (2003).
98. Cooper, G. M., Brudno, M., Green, E. D., Batzoglou, S. & Sidow, A. Quantitative estimates of sequence divergence for comparative analyses of mammalian genomes. *Genome Res.* **13**, 813–820 (2003).
99. Huckins, C. The spermatogonial stem cell population in adult rats. I. Their morphology, proliferation and maturation. *Anat. Rec.* **169**, 533–557 (1971).
100. Clermont, Y. Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal. *Physiol. Rev.* **52**, 198–236 (1972).
101. Makova, K. D., Yang, S. & Chiaromonte, F. Insertions and deletions are male biased too: A whole-genome analysis in rodents. *Genome Res.* **14**, 567–573 (2004).
102. Sundstrom, H., Webster, M. T. & Ellegren, H. Is the rate of insertion and deletion mutation male biased? Molecular evolutionary analysis of avian and primate sex chromosome sequences. *Genetics* **164**, 259–268 (2003).
103. Chang, B. H. & Li, W. H. Estimating the intensity of male-driven evolution in rodents by using X-linked and Y-linked Ube 1 genes and pseudogenes. *J. Mol. Evol.* **40**, 70–77 (1995).
104. Chang, B. H., Shimmin, L. C., Shyue, S. K., Hewett-Emmett, D. & Li, W. H. Weak male-driven molecular evolution in rodents. *Proc. Natl Acad. Sci. USA* **91**, 827–831 (1994).
105. Bird, A. P. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8**, 1499–1504 (1980).
106. Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA* **90**, 11995–11999 (1993).
107. Jensen-Seaman, M. I. *et al.* Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res.* **14**, 528–538 (2004).
108. Birdsall, J. A. Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Mol. Biol. Evol.* **19**, 1181–1197 (2002).
109. Montoya-Burgos, J. I., Boursot, P. & Galtier, N. Recombination explains isochores in mammalian genomes. *Trends Genet.* **19**, 128–130 (2003).
110. Bray, N. & Pachter, L. MAVID Constrained ancestral alignment of multiple sequence. *Genome Res.* **14**, 693–699 (2004).
111. Yap, V. B. & Pachter, L. Identification of evolutionary hotspots in the rodent genomes. *Genome Res.* **14**, 574–579 (2004).
112. Hubbard, T. *et al.* The Ensembl genome database project. *Nucleic Acids Res.* **30**, 38–41 (2002).
113. Vitt, U. *et al.* Identification of candidate disease genes by EST alignments, synteny and expression and verification of Ensembl genes on rat chromosome 1q43–54. *Genome Res.* **14**, 640–650 (2004).
114. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).
115. Guigo, R., Knudsen, S., Drake, N. & Smith, T. Prediction of gene structure. *J. Mol. Biol.* **226**, 141–157 (1992).
116. Solovye, V. V., Salamov, A. A. & Lawrence, C. B. Identification of human gene structure using linear discriminant functions and dynamic programming. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **3**, 367–375 (1995).
117. Parra, G. *et al.* Comparative gene prediction in human and mouse. *Genome Res.* **13**, 108–117 (2003).
118. Alexandersson, M., Cawley, S. & Pachter, L. SLAM—Cross-species gene finding with a generalized pair hidden Markov model. *Genome Res.* **13**, 496–502 (2003).
119. Flicke, P., Keibler, E., Hu, P., Korf, I. & Brent, M. R. Leveraging the mouse genome for gene prediction in human: from whole-genome shotgun reads to a global synteny map. *Genome Res.* **13**, 46–54 (2003).
120. Korf, I., Flicke, P., Duan, D. & Brent, M. R. Integrating genomic homology into gene structure prediction. *Bioinformatics* **17** (suppl. 1), S140–S148 (2001).
121. Wu, J. Q., Shteynberg, D., Arumugam, M., Gibbs, R. A. & Brent, M. R. Identification of rat genes by TWINSKAN gene prediction, RT–PCR, and direct sequencing. *Genome Res.* **14**, 655–671 (2004).
122. Dewey, C. *et al.* Accurate identification of novel human genes through simultaneous gene prediction in human, mouse and rat. *Genome Res.* **14**, 661–664 (2004).
123. Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
124. Puente, X. S. & Lopez-Otin, C. A genomic analysis of rat proteases and protease inhibitors. *Genome Res.* **14**, 609–622 (2004).
125. Puente, X. S., Sanchez, L. M., Overall, C. M. & Lopez-Otin, C. Human and mouse proteases: a comparative genomic approach. *Nature Rev. Genet.* **4**, 544–558 (2003).
126. Hurst, L. D. The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* **18**, 486–487 (2002).
127. Makalowski, W. & Boguski, M. S. Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proc. Natl Acad. Sci. USA* **95**, 9407–9412 (1998).
128. Wolfe, K. H. & Sharp, P. M. Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. *J. Mol. Evol.* **37**, 441–456 (1993).
129. Modrek, B. & Lee, C. J. Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nature Genet.* **34**, 177–180 (2003).
130. Nekrutenko, A., Makova, K. D. & Li, W. H. The K_A/K_S ratio test for assessing the protein-coding potential of genomic regions: an empirical and simulation study. *Genome Res.* **12**, 198–202 (2002).

131. Nekrutenko, A., Chung, W. Y. & Li, W. H. An evolutionary approach reveals a high protein-coding capacity of the human genome. *Trends Genet.* **19**, 306–310 (2003).
132. Taylor, M. S., Ponting, C. P. & Copley, R. R. Occurrence and consequences of coding sequence insertions and deletions in mammalian genomes. *Genome Res.* **14**, 555–566 (2004).
133. Green, H. & Wang, N. Codon reiteration and the evolution of proteins. *Proc. Natl Acad. Sci. USA* **91**, 4298–4302 (1994).
134. Levinson, G. & Gutman, G. A. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**, 203–221 (1987).
135. Alba, M. M. & Guigo, R. Comparative analysis of amino acid repeats in rodents and humans. *Genome Res.* **14**, 549–554 (2004).
136. Alba, M. M., Santibanez-Koref, M. F. & Hancock, J. M. Conservation of polyglutamine tract size between mice and humans depends on codon interruption. *Mol. Biol. Evol.* **16**, 1641–1644 (1999).
137. Burge, C. B., Padgett, R. A. & Sharp, P. A. Evolutionary fates and origins of U12-type introns. *Mol. Cell* **2**, 773–785 (1998).
138. Ohno, S. *Evolution by Gene Duplication* (Springer, Berlin, 1970).
139. Emes, R. D., Goodstadt, L., Winter, E. E. & Ponting, C. P. Comparison of the genomes of human and mouse lays the foundation of genome zoology. *Hum. Mol. Genet.* **12**, 701–709 (2003).
140. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
141. Prince, V. E. & Pickett, F. B. Splitting pairs: the diverging fates of duplicated genes. *Nature Rev. Genet.* **3**, 827–837 (2002).
142. Hughes, A. L. *Adaptive Evolution of Genes and Genomes* Ch. 7, 143–179 (Oxford Univ. Press, New York, 1999).
143. Tagle, D. A. *et al.* Embryonic epsilon and gamma globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *J. Mol. Biol.* **203**, 439–455 (1988).
144. Altschul, S. F. & Lipman, D. J. Protein database searches for multiple alignments. *Proc. Natl Acad. Sci. USA* **87**, 5509–5513 (1990).
145. Gumucio, D. L. *et al.* Phylogenetic footprinting reveals a nuclear protein which binds to silencer sequences in the human gamma and epsilon globin genes. *Mol. Cell. Biol.* **12**, 4919–4929 (1992).
146. Hardison, R. *et al.* Comparative analysis of the locus control region of the rabbit beta-like gene cluster: HS3 increases transient expression of an embryonic epsilon-globin gene. *Nucleic Acids Res.* **21**, 1265–1272 (1993).
147. Boffelli, D. *et al.* Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* **299**, 1391–1394 (2003).
148. Eltnitski, L. *et al.* Distinguishing regulatory DNA from neutral sites. *Genome Res.* **13**, 64–72 (2003).
149. Loots, G. G., Ovcharenko, I., Pachter, L., Dubchak, I. & Rubin, E. M. rVista for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Res.* **12**, 832–839 (2002).
150. Pennacchio, L. A. & Rubin, E. M. Genomic strategies to identify mammalian regulatory sequences. *Nature Rev. Genet.* **2**, 100–109 (2001).
151. Margulies, E. H., Blanchette, M., Haussler, D. & Green, E. Identification and characterization of multi-species conserved sequences. *Genome Res.* **13**, 2507–2518 (2003).
152. Kolbe, D. *et al.* Regulatory potential scores from genome-wide 3-way alignments of human, mouse and rat. *Genome Res.* **14**, 700–707 (2004).
153. Wingender, E. *et al.* The TRANSFAC system on gene expression regulation. *Nucleic Acids Res.* **29**, 281–283 (2001).
154. Trinklein, N. D., Aldred, S. J., Saldanha, A. J. & Myers, R. M. Identification and functional analysis of human transcriptional promoters. *Genome Res.* **13**, 308–312 (2003).
155. Philipsen, S., Pruzina, S. & Grosveld, F. The minimal requirements for activity in transgenic mice of hypersensitive site 3 of the beta globin locus control region. *EMBO J.* **12**, 1077–1085 (1993).
156. Reddy, P. M. & Shen, C. K. Protein-DNA interactions *in vivo* of an erythroid-specific, human beta-globin locus enhancer. *Proc. Natl Acad. Sci. USA* **88**, 8676–8680 (1991).
157. Strauss, E. C. & Orkin, S. H. *In vivo* protein-DNA interactions at hypersensitive site 3 of the human beta-globin locus control region. *Proc. Natl Acad. Sci. USA* **89**, 5809–5813 (1992).
158. Hillier, L. W. *et al.* The DNA sequence of human chromosome 7. *Nature* **424**, 157–164 (2003).
159. Torrents, D., Suyama, M. & Bork, P. A genome-wide survey of human pseudogenes. *Genome Res.* **13**, 2559–2567 (2003).
160. Zhang, Z., Harrison, P. & Gerstein, M. Identification and analysis of over 2000 ribosomal protein pseudogenes in the human genome. *Genome Res.* **12**, 1466–1482 (2002).
161. Mulder, N. J. *et al.* The InterPro Database 2003 brings increased coverage and new features. *Nucleic Acids Res.* **31**, 315–318 (2003).
162. Oh, B., Hwang, S. Y., Solter, D. & Knowles, B. B. Spindlin, a major maternal transcript expressed in the mouse during the transition from oocyte to embryo. *Development* **124**, 493–503 (1997).
163. Garcia-Meunier, P., Etienne-Julian, M., Fort, P., Piechaczyk, M. & Bonhomme, F. Concerted evolution in the GAPDH family of retrotransposed pseudogenes. *Mamm. Genome* **4**, 695–703 (1993).
164. Smit, A. F. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.* **9**, 657–663 (1999).
165. Prak, E. T. & Kazazian, H. H. Jr Mobile elements and the human genome. *Nature Rev. Genet.* **1**, 134–144 (2000).
166. Ostertag, E. M. & Kazazian, H. H. Jr Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* **35**, 501–538 (2001).
167. Weiner, A. M. SINEs and LINEs: the art of biting the hand that feeds you. *Curr. Opin. Cell Biol.* **14**, 343–350 (2002).
168. Martin, S. L. & Bushman, F. D. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell. Biol.* **21**, 467–475 (2001).
169. Hayward, B. E., Zavanelli, M. & Furano, A. V. Recombination creates novel L1 (LINE-1) elements in *Rattus norvegicus*. *Genetics* **146**, 641–654 (1997).
170. Dewannieux, M., Esnault, C. & Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nature Genet.* **35**, 41–48 (2003).
171. Quentin, Y. A master sequence related to a free left Alu monomer (FLAM) at the origin of the B1 family in rodent genomes. *Nucleic Acids Res.* **22**, 2222–2227 (1994).
172. Cantrell, M. A. *et al.* An ancient retrovirus-like element contains hot spots for SINE insertion. *Genetics* **158**, 769–777 (2001).
173. Rothenburg, S., Eiben, M., Koch-Nolte, F. & Haag, F. Independent integration of rodent identifier (ID) elements into orthologous sites of some RT6 alleles of *Rattus norvegicus* and *Rattus rattus*. *J. Mol. Evol.* **55**, 251–259 (2002).
174. Roy-Engel, A. M. *et al.* Non-traditional Alu evolution and primate genomic diversity. *J. Mol. Biol.* **316**, 1033–1040 (2002).
175. Salem, A. H., Kilroy, G. E., Watkins, W. S., Jorde, L. B. & Batzer, M. A. Recently integrated Alu elements and human genomic diversity. *Mol. Biol. Evol.* **20**, 1349–1361 (2003).
176. Salem, A. H. *et al.* Alu elements and hominid phylogenetics. *Proc. Natl Acad. Sci. USA* **100**, 12787–127891 (2003).
177. Smit, A. F. Identification of a new, abundant superfamily of mammalian LTR-transposons. *Nucleic Acids Res.* **21**, 1863–1872 (1993).
178. Benit, L. *et al.* Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the Fv1 restriction gene. *J. Virol.* **71**, 5652–5657 (1997).
179. Costas, J. Molecular characterization of the recent intragenomic spread of the murine endogenous retrovirus MuERV-L. *J. Mol. Evol.* **56**, 181–186 (2003).
180. Emes, R. D., Beatson, S. A., Ponting, C. P. & Goodstadt, L. Evolution and comparative genomics of odorant- and pheromone-associated genes in rodents. *Genome Res.* **14**, 591–602 (2004).
181. Young, J. M. *et al.* Different evolutionary processes shaped the mouse and human olfactory receptor gene families. *Hum. Mol. Genet.* **11**, 535–546 (2002).
182. Zhang, X. & Firestein, S. The olfactory receptor gene superfamily of the mouse. *Nature Neurosci.* **5**, 124–133 (2002).
183. Rouquier, S., Blancher, A. & Giorgi, D. The olfactory receptor gene repertoire in primates and mouse: evidence for reduction of the functional fraction in primates. *Proc. Natl Acad. Sci. USA* **97**, 2870–2874 (2000).
184. Clark, A. J., Hickman, J. & Bishop, J. A 45-kb DNA domain with two divergently orientated genes is the unit of organisation of the murine major urinary protein genes. *EMBO J.* **3**, 2055–2064 (1984).
185. Mural, R. J. *et al.* A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome. *Science* **296**, 1661–1671 (2002).
186. Cavagioni, A. & Mucignat-Caretta, C. Major urinary proteins, alpha2u-globulins and aphrodisin. *Biochim. Biophys. Acta* **1482**, 218–228 (2000).
187. Hurst, J. L. *et al.* Individual recognition in mice mediated by major urinary proteins. *Nature* **414**, 631–634 (2001).
188. Danielson, P. B. The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. *Curr. Drug Metab.* **3**, 561–597 (2002).
189. Nelson, D. R. Cytochrome P450 and the individuality of species. *Arch. Biochem. Biophys.* **369**, 1–10 (1999).
190. Scarborough, P. E., Ma, J., Qu, W. & Zeldin, D. C. P450 subfamily CYP2J and their role in the bioactivation of arachidonic acid in extrahepatic tissues. *Drug Metab. Rev.* **31**, 205–234 (1999).
191. Willson, T. M. & Klierer, S. A. PXR, CAR and drug metabolism. *Nature Rev. Drug Discov.* **1**, 259–266 (2002).
192. Gurates, B. *et al.* WT1 and DAX-1 inhibit aromatase P450 expression in human endometrial and endometriotic stromal cells. *J. Clin. Endocrinol. Metab.* **87**, 4369–4377 (2002).
193. Zhang, Z. *et al.* Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome. *Genome Res.* **14**, 580–590 (2004).
194. Lopez-Otin, C. & Overall, C. M. Protease degradomics: a new challenge for proteomics. *Nature Rev. Mol. Cell Biol.* **3**, 509–519 (2002).
195. Stenson, P. D. *et al.* Human Gene Mutation Database (HGMD): 2003 update. *Hum. Mutat.* **21**, 577–581 (2003).
196. Huang, H. *et al.* Evolutionary conservation of human disease gene orthologs in the rat and mouse genomes. *Genome Biol.* (submitted).
197. Duret, L. & Mouchiroud, D. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol. Biol. Evol.* **17**, 68–74 (2000).
198. Reddy, P. S. & Housman, D. E. The complex pathology of trinucleotide repeats. *Curr. Opin. Cell Biol.* **9**, 364–372 (1997).
199. The International HapMap Consortium. The International HapMap Project. *Nature* **426**, 789–796 (2003).
200. Wade, C. M. *et al.* The mosaic structure of variation in the laboratory mouse genome. *Nature* **420**, 574–578 (2002).
201. Zimdahl, H. *et al.* A SNP map of the rat genome generated from cDNA sequences. *Science* **303**, 807 (2004).
202. Mendell, J. T. & Dietz, H. C. When the message goes awry: disease-producing mutations that influence mRNA content and performance. *Cell* **107**, 411–414 (2001).
203. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
204. Marra, M. *et al.* A map for sequence analysis of the *Arabidopsis thaliana* genome. *Nature Genet.* **22**, 265–270 (1999).
205. Marra, M. A. *et al.* High throughput fingerprint analysis of large-insert clones. *Genome Res.* **7**, 1072–1084 (1997).
206. The International Human Genome Mapping Consortium. A physical map of the human genome. *Nature* **409**, 934–941 (2001).
207. Schein, J. E. A. in *Bacterial Artificial Chromosomes: Methods and Protocols* (eds Zhao, S. & Stodolsky, M.) 143–156 (Humana, Totowa, New Jersey, 2004).
208. Soderlund, C. I. *et al.* FPC: a system for building contigs from restriction fingerprinted clones. *Comput. Appl. Biosci.* **13**, 523–535 (1997).
209. Soderlund, C. S. *et al.* Contigs built with fingerprints, markers, and FPC V4.7. *Genome Res.* **10**, 1772–1787 (2000).
210. Ness, S. R. *et al.* Assembly of fingerprint contigs: parallelized FPC. *Bioinformatics* **18**, 484–485 (2002).
211. Woon, P. Y. *et al.* Construction and characterization of a 10-fold genome equivalent rat P1-derived artificial chromosome library. *Genomics* **50**, 306–316 (1998).
212. Watanabe, T. K. *et al.* A radiation hybrid map of the rat genome containing 5,255 markers. *Nature Genet.* **22**, 27–36 (1999).

213. Schuler, G. D. Sequence mapping by electronic PCR. *Genome Res.* **7**, 541–550 (1997).
214. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
215. Blanchette, M. *et al.* Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004).
216. Brudno, M. *et al.* LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **13**, 721–731 (2003).
217. Brudno, M. *et al.* Automated whole-genome multiple alignment of rat, mouse, and human. *Genome Res.* **14**, 685–692 (2004).
218. Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J. & Eichler, E. E. Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res.* **11**, 1005–1017 (2001).
219. Schwartz, S. *et al.* MultiPipMaker and supporting tools: Alignments and analysis of multiple genomic DNA sequences. *Nucleic Acids Res.* **31**, 3518–3524 (2003).
220. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
221. Birney, E. & Durbin, R. Using GeneWise in the *Drosophila* annotation experiment. *Genome Res.* **10**, 547–548 (2000).
222. Yang, Z., Goldman, N. & Friday, A. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. *Mol. Biol. Evol.* **11**, 316–324 (1994).
223. Chakrabarti, K. & Pachter, L. Visualization of multiple genome annotations and alignments with the K-BROWSER. *Genome Res.* **14**, 716–720 (2004).
224. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
225. Haldi, M. L. *et al.* Construction of a large-insert yeast artificial chromosome library of the rat genome. *Mamm. Genome* **8**, 284 (1997).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements Work at Baylor College of Medicine was supported by a grant from the NHGRI and NHLBI to R.A.G. Work at Genome Therapeutics was supported by grants from the NHGRI to D.S. A.S. acknowledges support from the NIGMS. M.B. acknowledges support from the NIH. N.H. was supported by the NGFN/BMBF (German Ministry for Research and Education). B.J.T. and J.M.Y. are supported by an NIH grant from the NIDCD. K.M.R. and G.M.C. are Howard Hughes Medical Institute Predoctoral Fellows. L.M.D'S., K.M. and K.J.K. are supported by training fellowships from the W. M. Keck Foundation to the Gulf Coast Consortia through the Keck Center for Computational and Structural Biology. Work at Case Western Reserve was supported in part by NIH grants to E.E.E. Work at IMIM was supported by a grant from Plan Nacional de I + D (Spain). M.M.A. acknowledges support from programme Ramón y Cajal and a grant from the Spanish Ministry of Science and Technology. Work at Universidad de Oviedo was supported by grants from the European Union, Obra Social Cajastur and Gobierno del Principado de Asturias. Work at Penn State University was supported by NHGRI grants. Work at the University of California Berkeley was supported by a grant from the NIH. Work at the Washington University School of Medicine Genome Sequencing Center and the British Columbia Cancer Agency Genome Sciences Centre was supported by an NIH grant. Work at UCSC and CHORI was supported by the NHGRI.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.A.G. (agibbs@bcm.tmc.edu). The genomic sequence is available under accession numbers AABR03000000 to AABR03137910 in the international sequence databases (GenBank, DDBJ and EMBL).

Rat Genome Sequencing Project Consortium (Participants are arranged under area of contribution, and then by institution.)

DNA sequencing: Baylor College of Medicine Richard A. Gibbs (Principal Investigator)¹, George M. Weinstock (Co-principal Investigator)¹, Michael L. Metzker¹, Donna M. Muzny¹, Erica J. Sodergren¹, Steven Scherer¹, Graham Scott¹, David Steffen¹, Kim C. Worley¹, Paula E. Burch¹, Geoffrey Okwuonu¹, Sandra Hines¹, Lora Lewis¹, Christine DeRamo¹, Oliver Delgado¹, Shannon Dugan-Rocha¹, George Miner¹, Margaret Morgan¹, Alicia Hawes¹, Rachel Gill¹, Celera Robert A. Holt (Principal Investigator)^{2,3}, Mark D. Adams^{3,4}, Peter G. Amanatides^{3,5}, Holly Baden-Tillson^{3,6}, Mary Barnstead^{3,7}, Soo Chin³, Cheryl A. Evans³, Steve Ferriera^{3,8}, Carl Foster³, Anna Glodek^{3,9}, Ziping Gu³, Don Jennings³, Cheryl L. Kraft^{3,10}, Trixie Nguyen³, Cynthia M. Pfannkoch^{3,6}, Cynthia Sitter^{3,11}, Granger G. Sutton³, J. Craig Venter^{3,8}, Trevor Woodage³; **Genome Therapeutics** Douglas Smith (Principal Investigator)^{12,13}, Hong-Mei Lee¹², Erik Gustafson^{12,13}, Patrick Cahill¹², Arnold Kana¹², Lynn Doucette-Stamm^{12,13}, Keith Weinstock¹², Kim Fechtel¹²; **University of Utah** Robert B. Weiss (Principal Investigator)¹⁴, Diane M. Dunn¹⁴; **NISC Comparative Sequencing Program, NHGRI** Eric D. Green¹⁵, Robert W. Blakesley¹⁵, Gerard G. Bouffard¹⁵

BAC library production: Children's Hospital Oakland Research Institute Pieter J. de Jong (Principal Investigator)¹⁶, Kazutoyo Osoegawa¹⁶, Baoli Zhu¹⁶

BAC fingerprinting: British Columbia Cancer Agency, Canada's Michael Smith Genome Sciences Centre Marco Marra (Principal Investigator)², Jacqueline Schein (Principal Investigator)², Ian Bosdet², Chris Fjell², Steven Jones², Martin Krzywinski², Carrie Mathewson², Asim Siddiqui², Natasja Wye²; **Genome Sequencing Center, Washington University School of Medicine** John McPherson^{1,17}

BAC end sequencing: TIGR Shaying Zhao (Principal Investigator)¹⁸, Claire M. Fraser¹⁸, Jyoti Shetty¹⁸, Sofiya Shatsman¹⁸, Keita Geer¹⁸, Yixin Chen¹⁸, Sofya Abramzon¹⁸, William C. Nierman¹⁸

Sequence assembly: Baylor College of Medicine Richard A. Gibbs (Principal Investigator)¹, George M. Weinstock (Principal Investigator)¹, Paul H. Havlak¹, Rui Chen¹, K. James Durbin¹, Amy Egan¹, Yanru Ren¹, Xing-Zhi Song¹, Bingshan Li¹, Yue Liu¹, Xiang Qin¹

Analysis and annotation: Affymetrix Simon Cawley¹⁹; **Baylor College of Medicine** George M. Weinstock (Coordinator)¹, Kim C. Worley (Overall Coordinator)¹, A. J. Cooney²⁰, Richard A. Gibbs¹, Lisa M. D'Souza¹, Kirt Martin¹, Jia Qian Wu¹, Manuel L. Gonzalez-Garay¹, Andrew R. Jackson¹, Kenneth J. Kalafus^{1,58}, Michael P. McLeod¹, Aleksandar Milosavljevic¹, Davinder Virk¹, Andrei Volkov¹, David A. Wheeler¹, Zhengdong Zhang¹; **Case Western Reserve University** Jeffrey A. Bailey⁴, Evan E. Eichler⁴, Eray Tuzun⁴; **EBI, Wellcome Trust Genome Campus** Ewan Birney²¹, Emmanuel Mongin²¹, Abel Ureta-Vidal²¹, Cara Woodwark²¹; **EMBL, Heidelberg** Evgeny Zdobnov²², Peer Bork^{22,23}, Mikita Suyama²², David Torrents²²; **Fraunhofer-Chalmers Research Centre for Industrial Mathematics, Gothenburg** Marina Alexandersson²⁴; **Fred Hutchinson Cancer Research Center** Barbara J. Trask²⁵, Janet M. Young²⁵; **Genome Therapeutics** Douglas Smith (Principal Investigator)^{12,13}, Hui Huang¹², Kim Fechtel¹², Huajun Wang¹², Heming Xing¹², Keith Weinstock¹²; **Incyte Corporation** Sue Daniels²⁶, Darryl Gietzen²⁶, Jeanette Schmidt²⁶, Kristian Stevens²⁶, Ursula Vitt²⁶, Jim Wingrove²⁶; **Institut Municipal d'Investigacio Medica, Barcelona** Francisco Camara²⁷, M. Mar Albà²⁷, Josep F. Abril²⁷, Roderic Guigo²⁷; **The Institute for Systems Biology** Arian Smit²⁸; **Lawrence Berkeley National Laboratory** Inna Dubchak^{29,30}, Edward M. Rubin^{29,30}, Olivier Couronne^{29,30}, Alexander Poliakov²⁹; **Max Delbrück Center for Molecular Medicine** Norbert Hübner²³, Detlev Ganten²³, Claudia Goesele^{23,31}, Oliver Hummel^{23,31}, Thomas Kreitler^{23,31}, Young-Ae Lee²³, Jan Monti²³, Herbert Schulz²³, Heike Zimdahl²³;

Max Planck Institute for Molecular Genetics, Berlin Heinz Himmelbauer³¹, Hans Lehrach³¹; **Medical College of Wisconsin** Howard J. Jacob (Principal Investigator)³², Susan Bromberg³³, Jo Gullings-Handley³², Michael I. Jensen-Seaman³², Anne E. Kwitek³², Jozef Lazar³², Dean Pasko³³, Peter J. Tonellato³², Simon Twigger³²; **MRC Functional Genetics Unit, University of Oxford** Chris P. Ponting (Leader, Genes and Proteins Analysis Group)³⁴, Jose M. Duarte³⁴, Stephen Rice³⁴, Leo Goodstadt³⁴, Scott A. Beatson³⁴, Richard D. Emes³⁴, Eitan E. Winter³⁴, Caleb Webber³⁴; **MWG-Biotech** Petra Brandt³⁵, Gerald Nyakatura³⁵; **Pennsylvania State University** Margaret Adetobi³⁶, Francesca Chiaromonte³⁶, Laura Elnitski³⁶, Pallavi Eswara³⁶, Ross C. Hardison³⁶, Minmei Hou³⁶, Diana Kolbe³⁶, Katerina Makova³⁶, Webb Miller³⁶, Anton Nekrutenko³⁶, Cathy Riemer³⁶, Scott Schwartz³⁶, James Taylor³⁶, Shan Yang³⁶, Yi Zhang³⁶; **Roche Genetics and Roche Center for Medical Genomics** Klaus Lindpaintner³⁷; **Sanger Institute** T. Dan Andrews³⁸, Mario Caccamo³⁸, Michele Clamp³⁸, Laura Clarke³⁸, Valerie Curwen³⁸, Richard Durbin³⁸, Eduardo Eyras³⁸, Stephen M. Searle³⁸; **Stanford University** Gregory M. Cooper (Co-Leader, Evolutionary Analysis Group)³⁹, Serafim Batzoglou⁴⁰, Michael Brudno⁴⁰, Arend Sidow³⁹, Eric A. Stone³⁹; **The Center for the Advancement of Genomics** J. Craig Venter^{3,8}; **University of Arizona** Bret A. Payseur⁴¹; **Université de Montréal** Guillaume Bourque⁴²; **Universidad de Oviedo** Carlos López-Otin⁴³, Xose S. Puente⁴³; **University of California, Berkeley** Kushal Chakrabarti⁴⁴, Sourav Chatterji⁴⁴, Colin Dewey⁴⁴, Lior Pachter⁴⁵, Nicolas Bray⁴⁵, Von Bing Yap⁴⁵, Anat Caspi⁴⁶; **University of California, San Diego** Glenn Tesler⁴⁷, Pavel A. Pevzner⁴⁸; **University of California, Santa Cruz** David Haussler (Co-Leader, Evolutionary Analysis Group)⁴⁹, Krishna M. Roskin⁵⁰, Robert Baertsch⁵⁰, Hiram Clawson⁵⁰, Terrence S. Furey⁵⁰, Angie S. Hinrichs⁵⁰, Donna Karolchik⁵⁰, William J. Kent⁵⁰, Kate R. Rosenbloom⁵⁰, Heather Trumbower⁵⁰, Matt Weirauch^{36,50}; **University of Wales College of Medicine** David N. Cooper⁵¹, Peter D. Stenson⁵¹; **University of Western Ontario** Bin Ma⁵²; **Washington University** Michael Brent⁵³, Manimozhayan Arumugam⁵³, David Shteynberg⁵³; **Wellcome Trust Centre for Human Genetics, University of Oxford** Richard R. Copley⁵⁴, Martin S. Taylor⁵⁴; **The Wistar Institute** Harold Riethman⁵⁵, Uma Mudunuri⁵⁵

Scientific management: Jane Peterson⁵⁶, Mark Guyer⁵⁶, Adam Felsenfeld⁵⁶, Susan Old⁵⁷, Stephen Mockrin⁵⁷ & Francis Collins⁵⁶

Affiliations for participants: 1, Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, MS BCM226, One Baylor Plaza, Houston, Texas 77030, USA (<http://www.hgsc.bcm.tmc.edu>); 2, British Columbia Cancer Agency, Canada's Michael Smith Genome Sciences Centre, 600 W 10th Avenue, Vancouver, British Columbia V5Z 4E6, Canada (<http://www.bccsc.ca>); 3, Celera, 45 West Gude Drive, Rockville, Maryland 20850, USA; 4, Department of Genetics and the Center for Computational Genomics, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106, USA; 5, DSM Pharmaceuticals Inc., 5900 NW Greenville Blvd, Greenville, North Carolina 27834, USA; 6, The Institute for Biological Energy Alternatives (IBEA), 1901 Research Blvd, Rockville, Maryland 20850, USA; 7, Intronn, Inc., 910 Clopper Road, South Building, Gaithersburg, Maryland 20878, USA; 8, The Center for the Advancement of Genomics (TCAG), 1901 Research Blvd, Suite 600, Rockville, Maryland 20850, USA; 9, Avalon Pharmaceuticals, 20358 Seneca Meadows Parkway, Germantown, Maryland 20876, USA; 10, Basic Immunology Branch, Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Diseases (NIAID), NIH, DHHS, 6610 Rockledge Blvd, Room 3005, Bethesda, Maryland 20892-7612, USA; 11, DynPort Vaccine Company, LLC, 64 Thomas Jefferson Drive, Frederick, Maryland 21702, USA; 12, Genome Therapeutics Corporation, 100 Beaver Street, Waltham, Massachusetts 02453, USA; 13, Agencourt Bioscience Corporation, 100 Cummings Center, Beverly, Massachusetts 01915, USA; 14, Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112, USA; 15, NIH Intramural Sequencing Center (NISC) and Genome Technology Branch, National Human Genome Research Institute (NHGRI), National Institutes of Health, Bethesda, Maryland 20892, USA; 16, BACPAC Resources, Children's Hospital Oakland Research Institute, 747 52nd Street, Oakland, California 94609, USA (<http://bacpac.chori.org>); 17, Genome Sequencing Centre, Washington University School of Medicine, 4444 Forest Park Blvd, St Louis, Missouri 63108, USA (<http://genome.wustl.edu>); 18, The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, Maryland 20850, USA (<http://www.tigr.org>); 19, Affymetrix, 6550 Valjejo St, Suite 100, Emeryville, California 94608, USA; 20, Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA; 21, EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK; 22, EMBL, Meyerhofstrasse 1, Heidelberg 69117, Germany; 23, Max Delbrück Center for Molecular Medicine (MDC), Experimental Genetics of Cardiovascular Disease, Robert-Rössle-Strasse 10, Berlin 13125, Germany (<http://www.mdc-berlin.de/ratgenomel>); 24, Fraunhofer-Chalmers Research Centre for Industrial Mathematics, Chalmers Science Park, S-412 88 Gothenburg, Sweden; 25, Division of Human Biology, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N., C3-168, Seattle, Washington 98109, USA (<http://www.fhcr.org/labs/trask/>); 26, Incyte Corporation, 3160 Porter Drive, Palo Alto, California 94304, USA (<http://www.incyte.com>); 27, Grup de Recerca en Informàtica Biomèdia, Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra, and Programa de Bioinformàtica i Genòmica, Centre de Regulació Genòmica, C/ Dr. Aiguader 80, 08003 Barcelona, Catalonia, Spain; 28, Computational Biology Group, The Institute for Systems Biology, 1441 North 34th Street, Seattle, Washington 98103, USA; 29, Genomics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, California 94720, USA (<http://www.lbl.gov>); 30, US Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, California 94598, USA (<http://jgi.doe.gov>); 31, Max Planck Institute for Molecular Genetics, Ihnestrasse 73, Berlin 14195, Germany; 32, Human and Molecular Genetics Center, Bioinformatics Research Center, and Department of Physiology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA; 33, Rat Genome Database, Bioinformatics Research Center, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA; 34, MRC Functional Genetics Unit, University of Oxford, Department of Human Anatomy and Genetics, South Parks Road, Oxford OX1 3QX, UK; 35, MWG-Biotech, Anzing Strasse 7a, Ebersberg 85560, Germany; 36, Center for Comparative Genomics and Bioinformatics, Huck Institutes of Life Sciences, Departments of Biology, Statistics, Biochemistry and Molecular Biology, Computer Science and Engineering, and Health Evaluation Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA; 37, Roche Genetics and Roche Center for Medical Genomics, F. Hoffmann-La Roche Ltd, 4070 Basel, Switzerland; 38, Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK; 39, Departments of Pathology and Genetics, Stanford University, Stanford, California 94305, USA; 40, S256 James H. Clark Center, Department of Computer Science, Stanford University, Stanford, California 94305, USA; 41, Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA; 42, Centre de Recherches Mathématiques, Université de Montréal, 2920 Chemin de la tour, Montréal, Quebec H3T 1J8, Canada (<http://www.crm.umontreal.ca>); 43, Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología, Universidad de Oviedo, 33006 Oviedo, Spain (<http://web.uniovi.es/degradome>); 44, Department of Electrical Engineering and Computer Science, University of California Berkeley, Berkeley, California 94720, USA; 45, Department of Mathematics, University of California Berkeley, Berkeley, California 94720, USA; 46, Bioengineering Graduate Group, University of California Berkeley, Berkeley, California 94720, USA; 47, University of California, San Diego, Department of Mathematics, 9500 Gilman Drive, San Diego, California 92093-0112, USA (<http://www-cse.ucsd.edu/groups/bioinformatics>); 48, University of California, San Diego, Department of Computer Science and Engineering, 9500 Gilman Drive, San Diego, California 92093-0114, USA (<http://www-cse.ucsd.edu/groups/bioinformatics>); 49, Howard Hughes Medical Institute, Center for Biomolecular Science & Engineering, Mailstop SOE, Baskin School of Engineering, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA; 50, UCSC Genome Bioinformatics Group, Center for Biomolecular Science and Engineering, Mailstop SOE, Baskin School of Engineering, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, USA; 51, Institute of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff, CF14 4XN, UK; 52, Department of Computer Science, University of Western Ontario, London, Ontario N6A 5B7, Canada; 53, Laboratory for Computational Genomics, Campus Box 1045, Washington University, St Louis, Missouri 63130, USA (<http://genes.cse.wustl.edu>); 54, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; 55, The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104, USA; 56, US National Institutes of Health, National Human Genome Research Institute, 31 Center Drive, Bethesda, Maryland 20892, USA; 57, US National Institutes of Health, National Heart, Lung, and Blood Institute, Bethesda, Maryland 20892, USA; 58, Program in Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

The DNA sequence and analysis of human chromosome 13

A. Dunham¹, L. H. Matthews¹, J. Burton¹, J. L. Ashurst¹, K. L. Howe¹, K. J. Ashcroft¹, D. M. Beare¹, D. C. Burford¹, S. E. Hunt¹, S. Griffiths-Jones¹, M. C. Jones¹, S. J. Keenan¹, K. Oliver¹, C. E. Scott¹, R. Ainscough¹, J. P. Almeida¹, K. D. Ambrose¹, D. T. Andrews¹, R. I. S. Ashwell¹, A. K. Babbage¹, C. L. Bagguley¹, J. Bailey¹, R. Bannerjee¹, K. F. Barlow¹, K. Bates¹, H. Beasley¹, C. P. Bird¹, S. Bray-Allen¹, A. J. Brown¹, J. Y. Brown¹, W. Burrill¹, C. Carder¹, N. P. Carter¹, J. C. Chapman¹, M. E. Clamp¹, S. Y. Clark¹, G. Clarke¹, C. M. Clee¹, S. C. M. Clegg¹, V. Cobley¹, J. E. Collins¹, N. Corby¹, G. J. Coville¹, P. Deloukas¹, P. Dhami¹, I. Dunham¹, M. Dunn¹, M. E. Earthrowl¹, A. G. Ellington¹, L. Faulkner¹, A. G. Frankish¹, J. Frankland¹, L. French¹, P. Garner¹, J. Garnett¹, J. G. R. Gilbert¹, C. J. Gilson¹, J. Ghorl¹, D. V. Grafham¹, S. M. Gribble¹, C. Griffiths¹, R. E. Hall¹, S. Hammond¹, J. L. Harley¹, E. A. Hart¹, P. D. Heath¹, P. J. Howden¹, E. J. Huckle¹, P. J. Hunt¹, A. R. Hunt¹, C. Johnson¹, D. Johnson¹, M. Kay¹, A. M. Kimberley¹, A. King¹, G. K. Laird¹, C. J. Langford¹, S. Lawlor¹, D. A. Leongamornlert¹, D. M. Lloyd¹, C. Lloyd¹, J. E. Loveland¹, J. Lovell¹, S. Martin¹, M. Mashreghi-Mohammadi¹, S. J. McLaren¹, A. McMurray¹, S. Milne¹, M. J. F. Moore¹, T. Nickerson¹, S. A. Palmer¹, A. V. Pearce¹, A. I. Peck¹, S. Pelan¹, B. Phillimore¹, K. M. Porter¹, C. M. Rice¹, S. Searle¹, H. K. Sehra¹, R. Shownkeen¹, C. D. Skuce¹, M. Smith¹, C. A. Steward¹, N. Sycamore¹, J. Tester¹, D. W. Thomas¹, A. Tracey¹, A. Tromans¹, B. Tubby¹, M. Wall¹, J. M. Wallis¹, A. P. West¹, S. L. Whitehead¹, D. L. Willey¹, L. Wilming¹, P. W. Wray¹, M. W. Wright², L. Young¹, A. Coulson¹, R. Durbin¹, T. Hubbard¹, J. E. Sulston¹, S. Beck¹, D. R. Bentley¹, J. Rogers¹ & M. T. Ross¹

¹The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

²HUGO Gene Nomenclature Committee, The Galton Laboratory, Wolfson House, 4 Stephenson Way, London, NW1 2HE, UK

Chromosome 13 is the largest acrocentric human chromosome. It carries genes involved in cancer including the breast cancer type 2 (*BRCA2*) and retinoblastoma (*RB1*) genes, is frequently rearranged in B-cell chronic lymphocytic leukaemia, and contains the *DAOA* locus associated with bipolar disorder and schizophrenia. We describe completion and analysis of 95.5 megabases (Mb) of sequence from chromosome 13, which contains 633 genes and 296 pseudogenes. We estimate that more than 95.4% of the protein-coding genes of this chromosome have been identified, on the basis of comparison with other vertebrate genome sequences. Additionally, 105 putative non-coding RNA genes were found. Chromosome 13 has one of the lowest gene densities (6.5 genes per Mb) among human chromosomes, and contains a central region of 38 Mb where the gene density drops to only 3.1 genes per Mb.

The draft sequence of the human genome¹ has provided the basis for a systematic effort to finish each chromosome^{2–5} in order to produce an accurate and detailed description of the entire genome. In common with the other acrocentric autosomes (14, 15, 21, and 22) the short arm of chromosome 13 is heterochromatic and contains families of repeated sequences, including the ribosomal RNA gene arrays⁶. The long arm is euchromatic and contains most or all of the protein-coding genes of the chromosome. We have completed the euchromatic sequence and examined the characteristics of this gene-poor autosome in relation to other human chromosomes, and to the corresponding sequence in other species. The annotation of the sequence is available via the Vertebrate Genome Annotation (VEGA) database (http://vega.sanger.ac.uk/Homo_sapiens), and will provide a platform for the continued study of medical genetics, genome instability and evolution of human chromosomes.

Clone map and finished sequence

The physical map of the long arm of chromosome 13 comprises five contigs. A minimally overlapping set of 863 clones (the 'tilepath') was selected that contained bacterial artificial chromosome (BAC) clones, supplemented with three yeast artificial chromosome (YAC) clones where no bacterial clones could be found (see Supplementary Table S1). The sequence derived from the tilepath clones comprises 95,564,076 base pairs (bp), determined to an accuracy well above 99.99% using procedures described previously⁷. At the proximal end, the finished sequence stretches into the pericentromeric region. Here the sequence is highly similar (93.9% sequence identity) to the pericentromeric region of chromosome 21. As a result, we have not so far been able to extend the map further towards the centromere. At the distal end, the finished sequence extends to within 15 kilobases

(kb) of the TTAGGG telomeric repeats (<http://www.wistar.upenn.edu/Riethman>). Six gaps remain in the tilepath despite our screening of genomic libraries containing a combined 87-fold coverage of the chromosome. All the gap sizes were estimated by fluorescence *in situ* hybridization (FISH) analysis and represent a combined total of 1,665 kb. The long arm of chromosome 13 therefore measures 97.2 Mb, and the finished sequence covers 98.3% of the total (see <http://www.sanger.ac.uk/HGP/Chr13> and Supplementary Table S2).

The finished sequence contains all markers previously mapped to chromosome 13 in the deCODE genetic map⁸, 98% of those in the Marshfield genetic map⁹ and 98% of those in the Genemap '99 radiation hybrid map¹⁰. Furthermore, there is excellent concordance between marker order on the maps and in the sequence. An additional check of the integrity of the finished sequence was obtained by examining the alignment of paired end-sequences of fosmid or BAC inserts in the finished sequence for evidence of deletions or mis-assemblies (D. Jaffe, personal communication). These alignments were consistent throughout the euchromatic sequence, with a single exception. One BAC (AL355611) was found to be ~20 kb longer than three fosmids aligned to that region. Further examination of the physical map in this region confirmed that the fingerprint patterns of the other BACs in the map are entirely consistent with the sequence of AL355611. We therefore conclude that there is a length variation between the DNAs represented in the BAC versus the fosmid libraries.

Global analysis of the sequence was performed to provide data on distribution of genes, repeats, G+C content, CpG islands, single nucleotide polymorphisms (SNPs), and recombination rates. The complete analysis of the sequence can be seen in Supplementary Fig. 1. Some specific features are highlighted below.

Table 1 Summary of the annotated genes on chromosome 13, excluding predicted ncRNA genes

Category	Number of genes	Total gene length (bp)	Mean gene length (bp)	Mean exon length (bp)	Mean exons per gene
Known genes	231	24,563,589	106,731	306	12.1
Novel genes	97	5,552,022	57,261	319	6.2
Novel transcripts	145	4,362,536	30,138	222	3.7
Sum of known and novel genes and novel transcripts	473	33,559,841	73,106	296	8.3
Putative genes	160	1,543,286	9,654	276	2.1
Sum of non-pseudogenes	633	34,617,102	57,068	295	6.7
Processed pseudogenes	268	252,294	946	946	1.0
Unprocessed pseudogenes	28	916,219	32,998	139	8.7
Sum of all genes	929	35,388,461	40,152	320	5.2

The rows in bold are running totals of the preceding categories.

Gene index

Gene structures were curated manually following analysis of the finished sequence by alignment to all publicly available expressed sequences and application of gene prediction algorithms. A total of 929 gene structures were identified and classified (according to the definitions described in Methods) as known genes (231), novel genes (97), novel transcripts (145), putative genes (160) and pseudogenes (296, of which 268 are processed and 28 are unprocessed) (see Table 1 and Supplementary Table S3). At the time of this analysis, one gene that was previously assigned to chromosome 13 (*RHOK*) was missing from the analysed sequence, and two other genes were only partially represented (*ATP4B* in AL4421245 and *RASA3* in AL161774). The missing parts of *ATP4B* and *RASA3* have since been identified in the sequences of two recently finished clones, BX537316 and BX537329, respectively. The sequence of BX537316, which lies adjacent to gap 4 in the chromosome map (Supplementary Table S2), also contains the first two exons of *RHOK*.

The largest gene on chromosome 13 spans 1,468,562 bp (*GPC5*). The largest exon is the single exon of the Spastic ataxia gene (*SACS*), which measures 12,865 bp. The average number of exons per gene is 5.19, ranging from single-exon genes, of which there are 44, to a gene containing 84 exons (*MYCBP2*). Predicted proteins were classified using the Interpro database (<http://www.ebi.ac.uk/interpro/>). The most populous protein families on chromosome 13 (Supplementary Table S4) broadly reflect those that are most populous in the genome as a whole¹. We searched for CpG islands 5 kb upstream and 1 kb downstream of each annotated gene (see Methods), and found that 67% of known and novel genes are associated with a CpG island, which is consistent with previous reports^{2,5,11}.

Owing to the absence of coding potential and the lack of primary sequence conservation, non-coding RNA (ncRNA) genes pose a significant challenge for computational prediction. Furthermore, the discrimination between genes and pseudogenes by computational methods has not been possible for most ncRNA genes. However, recent software advances and the development of a database of structural RNA alignments (Rfam)¹² now permit a more comprehensive search for ncRNA genes (see Supplementary Table S5). Chromosome 13 contains only five of the 616 transfer RNA genes that have been found in the human genome. In addition, two tRNA pseudogenes were found. Rfam analysis identifies 98 additional putative ncRNA genes on chromosome 13, including 37 spliceosomal RNAs and 20 Y RNAs (components of the ribonucleoprotein particle Ro). Two of these genes, a U6 and a U2 spliceosomal RNA, are confirmed on the basis of near identity to previously published sequences.

MicroRNAs (miRNAs) are approximately 21-bp double-stranded products excised from a hairpin precursor, which regulate the expression of other genes by complementary binding to untranslated regions of a messenger RNA target. The miRNA Registry (<http://www.sanger.ac.uk/Software/Rfam/mirna/>) lists 147 identified human miRNAs, eight of which are found in two clusters in the chromosome 13 sequence. miR-15 and miR-16 locate

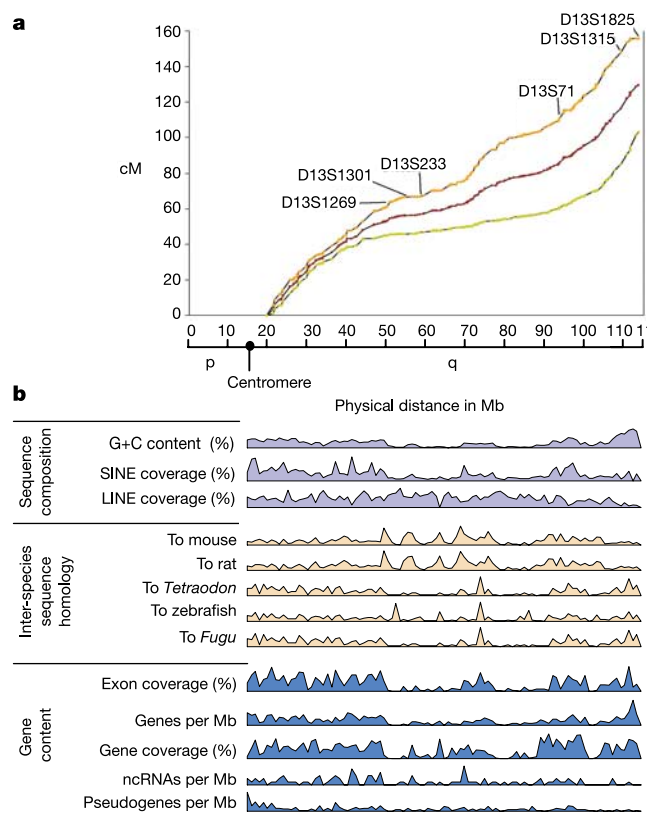
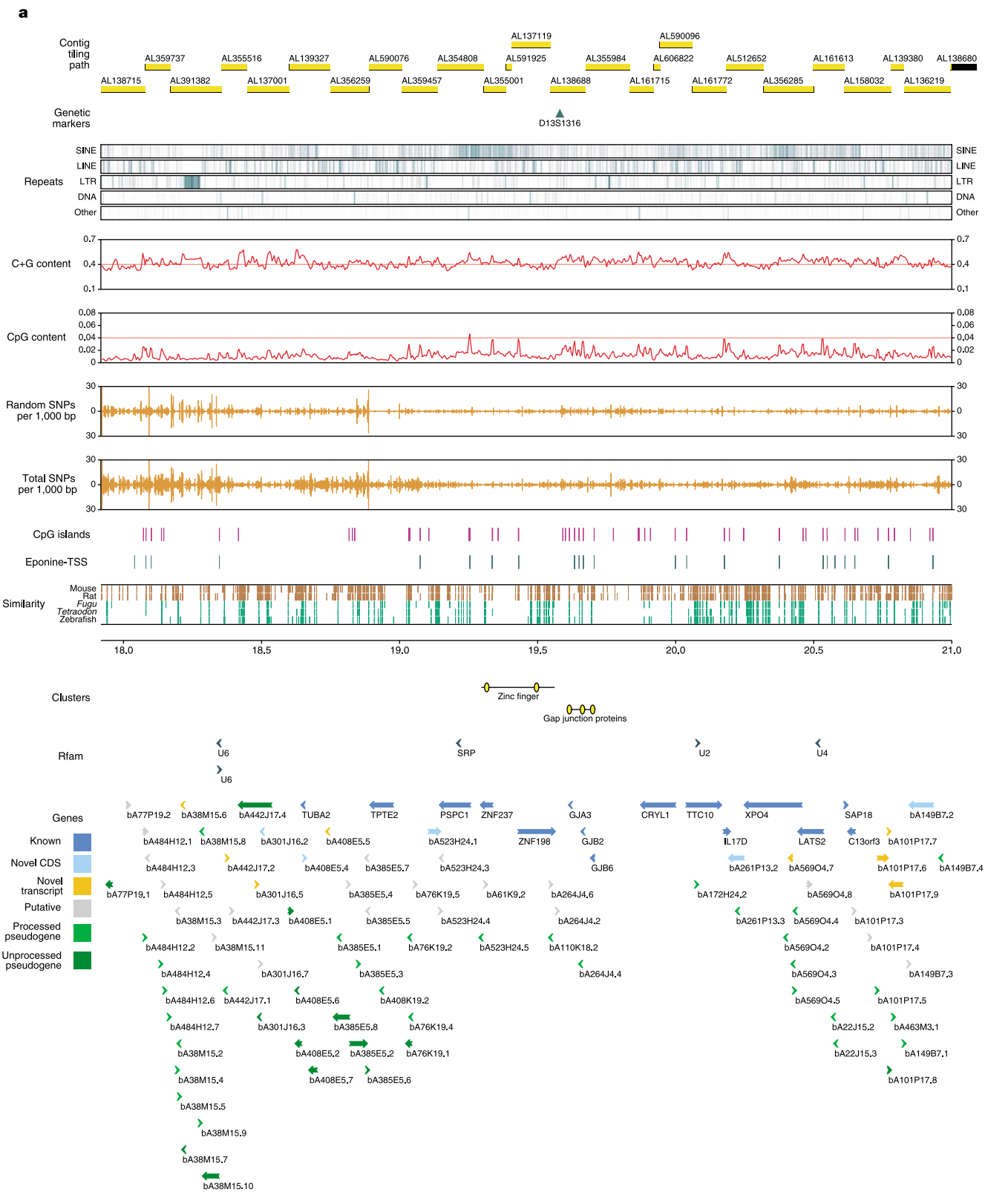


Figure 1 Genetic and physical characteristics of the chromosome 13 sequence. **a**, A comparison of physical and genetic distance along chromosome 13. Markers from the deCODE genetic map⁹ were localized in the sequence. Their locations in the genetic map are plotted on the y axis and locations in the sequence are on the x axis (the sequence starts at position 17,918,001 to allow for the chromosome short arm and centromere). The female genetic map is shown in orange, the male map in green and the sex average map in red. The loci shown mark the extent of the region of low gene density (D13S1269–D13S71), the recombination jungle (D13S1315–D13S1825) and the recombination desert (D13S1301–D13S233). **b**, Variation in features along the chromosome. The sequence was divided into 1-Mb non-overlapping sections, and each section was studied for the features shown. The data for each feature were normalized to set the largest figure at 100% and the lowest at 0%. For the repeat coverage, the interspecies sequence homologies (defined in the Methods), and the exon and gene coverage, the percentage of each 1-Mb window is calculated. The range of data points for each plot are as follows: % G+C content (33.5–52.4); % SINE coverage (3.5–25); % LINE coverage (9.2–31.4); % with homology to mouse (1.1–14.2); % with homology to rat (0.9–11.9); % with homology to *Tetraodon* (0–1.8); % with homology to zebrafish (0–3); % with homology to *Fugu* (0–1.9); % exon coverage (0–4.1); genes per Mb (0–38); % gene coverage (0–100); ncRNAs per Mb (0–6); pseudogenes per Mb (0–24).

to 13q14 and are separated by only 100 bp. These genes are deleted or downregulated in patients with chronic lymphocytic leukaemia¹³. A second group of six miRNAs is clustered within an 800 bp region of 13q32. The clustering of these miRNAs suggests that each set may be processed from the same primary transcript¹⁴. Both miRNAs in the first cluster and five of the genes in the second cluster are conserved in order and orientation on mouse chromosome 14.

Chromosome landscape

Chromosome 13 shows striking features of low gene density compared to the other finished, annotated autosomes, and also a very variable gene distribution. The overall gene density is the lowest of all the sequenced autosomes (see Table 2), with an average (excluding pseudogenes and ncRNA genes) of 6.5 genes per Mb (refs 2–5,15,16). This analysis extends and confirms previous



observations^{10,17,18}. Consistent with the low gene density, the G+C content of 38.5% and the predicted CpG island density of 5.4 Mb⁻¹ are considerably below the genome averages (Table 2). Exon coverage of chromosome 13 sequence is also substantially lower (1.3%) than that of other autosomes, except chromosome 7. However, the average gene length on chromosome 13 is 57 kb, which is almost

double that of other chromosomes (31 and 26 kb, for chromosomes 6 and 22, respectively). As a result, the 633 genes cover 37% of the sequence, which is only slightly lower than that of the other finished, annotated autosomes (Table 2).

Gene density varies considerably along the chromosome, as do other characteristics of the gene-rich and gene-poor regions (see

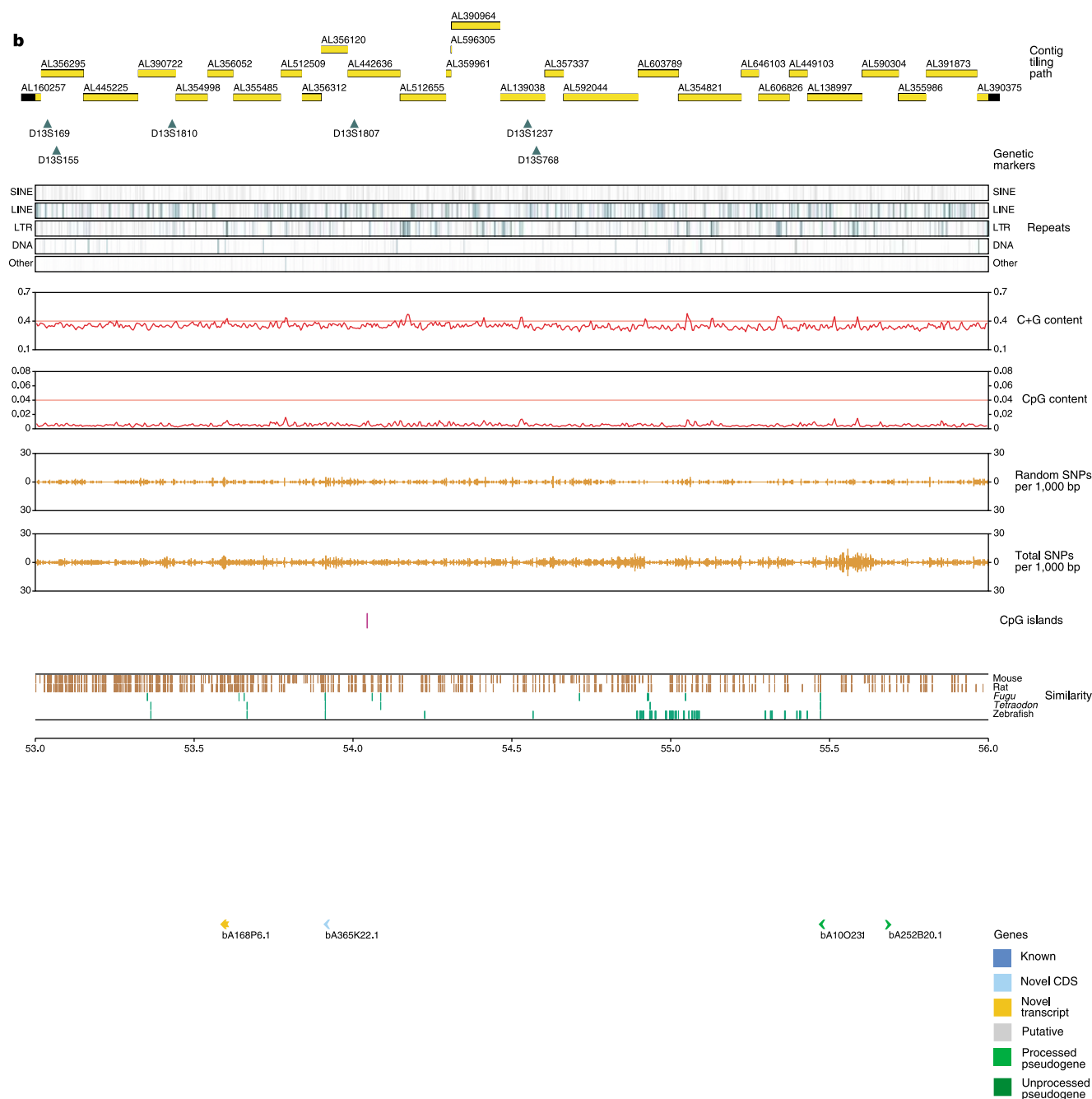


Figure 2 Characteristics of a gene-rich (a) and a gene-poor (b) region. The overlapping tilepath, labelled by accession number, is shown in yellow. Genetic markers from the deCODE map have been positioned on the sequence. Occurrences of repeats are shown as vertical turquoise bars. G+C content and CpG dinucleotide content are shown in overlapping windows of 8 kb, with adjacent windows overlapping by 4 kb. CpG islands are predicted using a modification of the CpG program developed by G. Micklem (personal communication). The positions of transcription start sites were predicted by the Eponine⁴⁹ program. Regions conserved in mouse and rat are shown by orange bars, and those conserved in *Fugu*, *Tetraodon* and zebrafish are shown by green bars. The Rfam

track contains the predicted ncRNA genes. Annotated gene structures are shown, subdivided into categories by colour. The direction of transcription is indicated by the arrow. The SNP tracks indicate the number of SNPs per kb. The random SNPs were generated by the whole-genome shotgun approach and were obtained from dbSNP by querying for SNPs produced by The SNP Consortium⁵⁰. The locations of clusters of two or more related genes within 1 Mb of each other are indicated. The scale shows the approximate Mb position along the chromosome. The chromosome view is available in Supplementary Fig. S1.

Fig. 1 and Supplementary Fig. 1). Here we define ‘gene-rich’ as containing more than 15 genes per Mb, and ‘gene-poor’ as containing fewer than 5 genes per Mb. A detailed picture of the characteristics of an example of each regional class is shown in Fig. 2. A comparison of the two regions can be seen in Table 3. There is a 37.8-Mb region, from 52.9–90.7 Mb, where the average gene density drops to 3 genes per Mb. This region actually comprises two gene-poor segments (52.9–71.9 Mb and 78.9–89.9 Mb) flanking a section with a gene density of 7 Mb⁻¹ (Fig. 1). The first gene-poor section contains a 3-Mb region with no genes (53.9–56.9 Mb).

The two major gene-rich areas are at either end of the q arm, with 90 of the predicted coding genes lying within 3 Mb of either end of the euchromatic region. As observed with other completed human chromosomes, gene-poor regions have a low G+C content, a low SINE coverage and a high LINE coverage, relative to genome averages. These trends are reversed for the gene-rich regions. Figure 1 shows that between positions 90 and 100 Mb, although there are very few genes, a large percentage of the region is covered by transcribed gene structures (exons plus introns). This region contains the largest gene on the chromosome (*GPC5*) as well as two others (*GPC6* and *HS6ST3*) each covering over 500 kb. In sharp contrast to the protein-coding genes, the ncRNA genes are distributed evenly between the gene-rich and gene-poor regions (see Fig. 1).

A total of 96,894 SNPs, from the dbSNP database, were mapped onto the sequence. The coding regions of the annotated genes contain 654 SNPs (1 SNP per 1.6 kb). These can be subdivided into 345 synonymous and 309 non-synonymous cSNPs. To analyse the overall distribution along the chromosome, a subset (38,069) of SNPs identified previously by alignment of random shotgun sequence to the draft sequence were plotted separately (see Fig. 2a, b and Supplementary Fig. 1). From this distribution plot, there is no obvious difference in the variation rate between the gene-rich and gene-poor areas. There is one region between positions 18.0 and 18.4 Mb (Figs 1 and 2a), where the SNP density is substantially higher than the average, reaching one SNP per 0.3 kb (1,329 SNPs in 400 kb). There is a known duplication with chromosome 21 in this region, and it is possible that the apparently high SNP density is due to the presence of paralogous sequence variants, as has been suggested previously¹⁶.

Around 5% of the human genome may be accounted for by segmental duplications, and this may play an important role in

genetic disease and genome evolution^{19,20}. A study by Cheung and colleagues¹⁹, which classified regions as segmental duplications if they show at least 90% homology over a minimum of 5 kb, suggested that there is approximately 1.8 Mb of duplicated sequence on chromosome 13. This sequence comprises 0.9 Mb of intrachromosomal and 1.2 Mb of interchromosomal duplications²⁰. This includes 0.3 Mb of sequence common to both categories. For example, the *TPIP* gene on chromosome 13 has undergone both inter- and intrachromosomal duplications. Guipponi and colleagues²¹ described phylogenetic analysis of the *TPTE* gene family, of which *TPIP* is a member, and suggested that all family members originate from a common ancestor since the divergence of human and mouse lineages, because there is only a single *Tpte* gene in mouse. There are seven other genes surrounding the mouse *Tpte* gene on chromosome 8, each of which has a functional homologue on human chromosome 13. This observation and the fact that four of the genes have homology only to chromosome 13 suggest that the human orthologue of the *Tpte* gene lies on chromosome 13. There have been a number of duplication events resulting in one functional copy and four pseudogenes of *TPIP* on chromosome 13. In addition, there is another functional member of the gene family, *TPTE*, on chromosome 21 and a number of pseudogenes on chromosomes 3, 15, 22 and Y.

Comparison to the genetic map

Recombination rates in the human genome are higher on average in females than males, and vary considerably along each chromosome. The sex-average genetic length of 13q is 129.52 cM, equating to an average recombination rate of 1.3 cM Mb⁻¹ (slightly higher than the genome average of 1.13 cM Mb⁻¹). Overall female and male rates for chromosome 13 are 1.6 and 1.1 cM Mb⁻¹ respectively. Figure 1 illustrates male, female and sex-averaged recombination as a function of chromosome position, correlated with a range of other features of the sequence annotation described above. In the 2 Mb closest to the centromere, the recombination rate in females compared to males is high (4.4 versus 1.7 cM Mb⁻¹) (Supplementary Table S6). Near the telomere, by contrast, the male recombination rate is the higher of the two (4.5 compared to 1.9 cM Mb⁻¹ in females).

Recombination rate appears to be correlated with gene density (Fig. 1). The recombination rate in male meiosis is particularly low (0.36 cM Mb⁻¹) in the central 42 Mb portion of the chromosome

Table 2 Comparison of chromosome 13 features with those of other sequenced autosomes

Category	Chromosome							
	6	7	13	14	20	21	22	Genome
% G+C content	40.0	41.0	38.5	40.9	44.1	40.9	47.8	41.0
% Repeat content	43.9	45.0	42.3	46.2	42	40.1	41.9	44.8
CpG island density (Mb ⁻¹)	13.1	9.5	5.4	20.2	11.1	-	16.5	10
% Gene coverage	42.2	36.5	37	43.6	42.4	31.0	50.0	-
% Exon coverage	2.2	1.4	1.3	2.3	2.4	-	5.0	-
Gene density (Mb ⁻¹)	9.2	7.5	6.5	10	12.2	6.7	16.3	~10
Sequence length (Mb)	166.8	153.8	95.5	87.4	59.2	33.5	33.5	2,862.7

The sequence length for the whole genome was derived from the Ensembl Genome Browser (http://www.ensembl.org/Homo_Sapiens). Other data were taken from the following references: chromosome 6 (ref. 5), chromosome 7 (ref. 4), chromosome 14 (ref. 3), chromosome 20 (ref. 2), chromosome 21 (ref. 16), chromosome 22 (ref. 11), genome (ref. 1).

Table 3 Comparison of the gene-rich (17.9–21 Mb) and gene-poor (53–56 Mb) regions

Feature	Gene-rich region	Gene-poor region
Gene density	16.5 genes Mb ⁻¹	0.7 genes Mb ⁻¹
G+C content	42.1%	34.9%
CpG island density	16.4 Mb ⁻¹	0.3 Mb ⁻¹
SINE coverage	19.5%	5.7%
LINE coverage	19.6%	25.2%

See Fig. 2a for gene-rich regions and Fig. 2b for gene-poor regions.

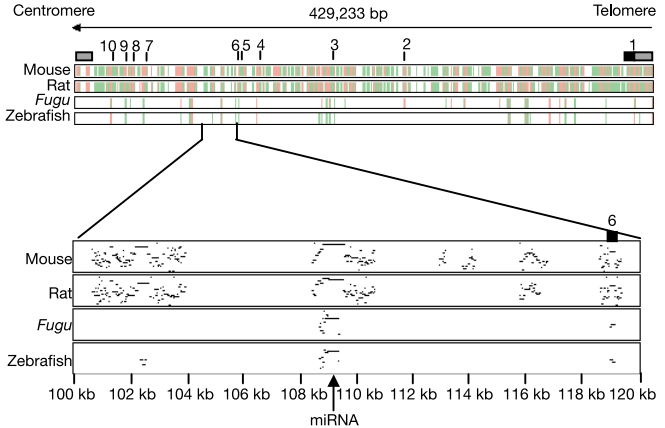


Figure 3 MultiPipMaker analysis²⁵ of the human *DACH1* gene showing conserved regions in mouse, rat, *Fugu* and zebrafish. **a**, The complete gene, with strongly aligned regions (at least 100 bp without a gap and with at least 70% nucleotide identity) in red. The green lines indicate regions conserved by local alignments. Exons are numbered. **b**, A more detailed picture of the area around exon 6. The arrow indicates a conserved intronic sequence which has 100% homology to a putative miRNA predicted using MiRscan²⁶.

between D13S1269 and D13S71, which corresponds to the region of lowest gene density on the chromosome. The female rate is 1.1 cM Mb^{-1} over the same region. Within this region there is also a noticeable increase in female recombination rate in the region of higher gene density (at position 72–80 Mb).

From analysis of the draft human genome sequence, Yu²² predicted two recombination 'deserts' on chromosome 13, which have a sex-average recombination rate less than or equal to 0.3 cM Mb^{-1} for physical distances up to 5 Mb in length, D13S164–D13S1228 (0.12 cM Mb^{-1}) and D13S1301–D13S233 (0.22 cM Mb^{-1}). Re-analysis of these regions using the finished sequence indicates that the former has a sex-average recombination rate of 0.84 cM Mb^{-1} , and so cannot be classified as a recombination desert. The latter, however, was contained within a 3.7-Mb region with a recombination rate of 0.16 cM Mb^{-1} , and so meets the criteria for a recombination desert. In the same study²², recombination 'jungles' were previously defined as having a sex-average recombination rate of 3 cM Mb^{-1} or more in regions up to 5 Mb, but none were predicted on chromosome 13. In our analysis there is one region of 4.1 Mb that meets the criteria for a jungle of high recombination, D13S1315–D13S1825 (3.2 cM Mb^{-1}).

Comparative analysis

The availability of assembled genome sequences for *Mus musculus* (mouse), *Rattus norvegicus* (rat), *Tetraodon nigroviridis* (*Tetraodon*), *Fugu rubripes* (*Fugu*) and *Danio rerio* (zebrafish), allowed us to identify regions where there is conservation between them and human chromosome 13 (see Supplementary Table S7). In general, greater similarity is expected between human and mouse or rat compared to *Fugu*, *Tetraodon* and zebrafish. We observed that 96% of known and novel genes on human chromosome 13 have exons conserved in both rodents, whereas only 81% have exons conserved in all three fishes. Most (95%) of the regions conserved in humans and fish correspond to annotated exons. By contrast, only 25% of the regions conserved in both mouse and rat overlap annotated exons, and at least some of the remainder are expected to result from selection for functionally important non-coding regions. It is of interest that regions with the largest coverage of conserved sequence between humans and rodents are found in gene-poor areas of the chromosome (see mouse and rat homology tracks in Fig. 1).

To estimate the completeness of our annotation of exons in protein-coding genes, we adopted the approach of previous studies^{2,5}. 2,553 regions on chromosome 13 were found to be conserved in all six reference genomes analysed here. Of these, 2,441 share overlap with 2,337 annotated exons on 13q, while the remaining 112 do not. On this basis, at least 95.4% ($2,337/2,337 + 112$) of exons on chromosome 13 are included in our annotation.

There are 112 conserved regions that do not correspond to annotated exons. Some or all of them might be exons that remain unconfirmed owing to a lack of transcriptional evidence, or they might be conserved regulatory regions. Recent reports suggest that non-exonic sequences conserved across species may be regulatory or structural elements^{23,24}. Nobrega and colleagues showed that seven evolutionarily conserved sequences, from in and around the mouse *Dach* gene, enhanced expression of reporter genes in specific tissues²³. These regions were also identified in our analysis. In fact, 28 of the 112 conserved regions are located in the introns of the *DACH1* gene. The MultiPipMaker program²⁵ was used to compare sequence from the *DACH1* gene in human, mouse, rat, *Fugu* and zebrafish (Fig. 3). One of the regions of conservation has a 100% match to a putative miRNA²⁶. Further studies are required to discover whether other of the conserved intronic regions have any regulatory effect.

Medical implications

Forty-eight mendelian conditions listed in Online Mendelian Inheritance in Man (OMIM) have been linked to genes on chromo-

some 13 (Supplementary Table S8). Thirty-five of these genes have been cloned and can be positioned on the sequence. The map and sequence data for chromosome 13 have contributed to the identification of a number of these. Notably, an early phase of the sequencing led to characterization of the breast cancer type-2 gene *BRCA2* (ref. 27). More recently, sequence from the 13q34 region was used in the identification of the *DAOA* locus, which is associated with schizophrenia and bipolar disorder^{28,29}.

The sequence is being applied to search for genes that are implicated in human disease but are as yet unidentified. B-cell chronic lymphocytic leukaemia (CLL) is one of the most common leukaemias in the western world and approximately 10% of CLL patients have a homozygous deletion in 13q14.3 (ref. 30): the chromosome 13 data have allowed further refinement of the region of minimal deletion³⁰. Follicular lymphoma (FL) accounts for approximately 25% of non-Hodgkin's lymphomas in the western world³¹. Using clones from the 13q32–q33 region of the tilepath, a region of recurrent amplification in FL patients has been identified³². Asthma and some forms of dermatitis are atopic or immunoglobulin E (IgE)-mediated diseases, which have shown consistent linkage to 13q14 (refs 33, 34). Zhang and colleagues have localized a quantitative trait locus influencing IgE levels and asthma using a dense SNP map of the region³⁵. In the same study, an association with severe clinical asthma was found with several alleles of the *PHF11* gene.

In addition to *BRCA2*, a number of other genes involved in tumorigenesis have been identified on chromosome 13, including the retinoblastoma gene *RB1* and the alveolar rhabdomyosarcoma gene *FOXO1A*. As described in Supplementary Table S8, several other such genes have been linked to chromosome 13 but remain to be identified. In the Mitleman database of recurrent chromosome aberrations in cancer (<http://cgap.nci.nih.gov/Chromosomes/Mitleman>), more than 1,400 cases affect chromosome 13. The availability of a tilepath of sequence clones will allow techniques such as array comparative genomic hybridization to refine the regions of rearrangement in these cases. □

Methods

Using the Genemap '99 radiation hybrid map¹⁰ as a starting point, a landmark mapping approach was used to seed and anchor BAC clone contigs³⁶. A total of 863 markers ($8.8 \text{ markers Mb}^{-1}$) were used to screen the RPCL-11 BAC library³⁷, resulting in the identification of 3,354 clones. Using the whole genome fingerprint data from the International Human Genome Mapping Consortium³⁸, the BACs were built into sequence-ready contigs using a combination of shared *HindIII* fingerprint bands and marker content. Gaps in the map were closed by directed screening of BAC, P1-derived artificial chromosome (PAC) and yeast artificial chromosome (YAC) libraries with probes derived from contig ends. Additionally, end-sequence data from the WIBR whole-genomic fosmid library were used in an effort to identify further contig extensions. All the gaps were sized using fluorescent *in situ* hybridization on DNA fibres, apart from gap 1 (Supplementary Table S2), which was sized on metaphase chromosomes.

All the tilepath clones processed at the Sanger Institute were sequenced using a shotgun approach and assembled as previously reported⁵. All finished clones meet or exceed the agreed international finishing standards of 99.99% sequence accuracy.

The finished genomic sequence was analysed using an automatic ENSEMBL³⁹ pipeline with modifications to aid the manual curation process. The G+C content of each clone sequence was analysed and putative CpG islands marked. CpG islands were predicted using a modification of the CpG program (G. Micklem, personal communication). The identification of interspersed repeats using RepeatMasker; simple repeats using Tandem Repeat Finder⁴⁰; matches to vertebrate complementary DNAs and expressed sequence tags (ESTs) using WU-BLASTN (W. Gish 1996–2002, <http://blast.wustl.edu>) and EST_GENOME; and *ab initio* gene prediction using GENESH and GENSCAN, were as described previously². A protein database combining non-redundant data from SwissProt⁴¹ and TrEMBL⁴¹ was searched using WU-BLASTX. ENSEMBL gene predictions, including the EST gene build, were displayed on each clone present in the finished sequence assembly. The predicted gene structures were manually annotated according to the human annotation workshop (HAWK) guidelines (<http://www.sanger.ac.uk/HGP/havana/hawk.shtml>). The gene categories used were as described on the VEGA website. Known genes are identical to known human cDNAs or protein sequences and should have an entry in Locuslink (<http://www.ncbi.nlm.nih.gov/LocusLink>). Novel genes have an open reading frame (ORF), are identical to spliced ESTs or have some similarity to other genes or proteins. Novel transcripts are similar to novel genes, but the ORF cannot be determined with confidence. Putative genes are identical to spliced human ESTs, but do not contain an ORF. Processed pseudogenes are non-functional copies of genes that lack

introns. Unprocessed pseudogenes are non-functional copies of genes that contain introns. Where possible, gene symbols were approved by the HUGO Gene Nomenclature Committee⁴².

tRNA genes were predicted using tRNAscan-SE v1.23 (ref. 43) in eukaryotic mode with default parameters and a threshold of 20 bits. Other ncRNAs were predicted by searching the chromosome sequence against the Rfam database of RNA families (version 4.1)⁴². Predicted ncRNA genes were compared by WU-BLASTN to reference sequences collected from the EMBL nucleotide sequence database⁴⁴, in an effort to discriminate between genes and pseudogenes. True ncRNA genes were operationally defined as BLAST hits with at least 95% sequence identity over 95% of the query length, as previously described¹. Published human microRNA gene sequences were downloaded from the miRNA Registry v2.0 and mapped to the chromosome sequence using BLASTN.

All markers from the deCODE genetic map were mapped back onto the finished sequence using a combination of electronic PCR⁴⁵ and SSAHA⁴⁶. For the comparative analysis with other genomes the following methods were used. The mouse and rat sequence were compared to the human sequence using BLASTZ⁴⁷. The programs axtBest and subsetAxt (W. J. Kent, <http://www.soe.ucsc.edu/~kent/src>) were used to post-process the resulting matches, as previously described⁴⁷, to select the best match and make the alignments relatively specific to exons by using a specific scoring matrix and threshold. The sequences from *Tetraodon*, *Fugu* and zebrafish were aligned to the chromosome using WU-TBLASTX, using the same scoring matrix, parameters and filtering strategy used in the Exofish procedure⁴⁸. The resources used for these comparisons are shown in the Supplementary Information.

Received 5 December 2003; accepted 27 January 2004; doi:10.1038/nature02379.

- International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Deloukas, P. *et al.* The DNA sequence and comparative analysis of human chromosome 20. *Nature* **414**, 865–871 (2001).
- Heilig, R. *et al.* The DNA sequence and analysis of human chromosome 14. *Nature* **421**, 601–607 (2003).
- Hillier, L. W. *et al.* The DNA sequence of human chromosome 7. *Nature* **424**, 157–164 (2003).
- Mungall, A. J. *et al.* The DNA sequence and analysis of human chromosome 6. *Nature* **425**, 805–811 (2003).
- Kehrer-Sawatzki, H. *et al.* Mapping of members of the low-copy-number repetitive DNA sequence family chAB4 within the p arms of human acrocentric chromosomes: characterization of Robertsonian translocations. *Chromosome Res.* **6**, 429–435 (1998).
- Felsenfeld, A., Peterson, J., Schloss, J. & Guyer, M. Assessing the quality of the DNA sequence from the Human Genome Project. *Genome Res.* **9**, 1–4 (1999).
- Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nature Genet.* **31**, 241–247 (2002).
- Broman, K. W., Murray, J. C., Sheffield, V. C., White, R. L. & Weber, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63**, 861–869 (1998).
- Deloukas, P. *et al.* A physical map of 30,000 human genes. *Science* **282**, 744–746 (1998).
- Dunham, I. *et al.* The DNA sequence of human chromosome 22. *Nature* **402**, 489–495 (1999).
- Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. & Eddy, S. R. Rfam: an RNA family database. *Nucleic Acids Res.* **31**, 439–441 (2003).
- Calin, G. A. *et al.* Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl Acad. Sci. USA* **99**, 15524–15529 (2002).
- Lee, Y., Jeon, K., Lee, J. T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663–4670 (2002).
- Collins, J. E. *et al.* Reevaluating human gene annotation: a second-generation analysis of chromosome 22. *Genome Res.* **13**, 27–36 (2003).
- Hattori, M. *et al.* The DNA sequence of human chromosome 21. *Nature* **405**, 311–319 (2000).
- Saccone, S. *et al.* Correlations between isochores and chromosomal bands in the human genome. *Proc. Natl Acad. Sci. USA* **90**, 11929–11933 (1993).
- Craig, J. M. & Bickmore, W. A. The distribution of CpG islands in mammalian chromosomes. *Nature Genet.* **7**, 376–382 (1994).
- Cheung, J. *et al.* Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biol.* [online] **4**, R25 (2003).
- Bailey, J. A. *et al.* Recent segmental duplications in the human genome. *Science* **297**, 1003–1007 (2002).
- Guipponi, M. *et al.* The murine orthologue of the Golgi-localized TPTE protein provides clues to the evolutionary history of the human TPTE gene family. *Hum. Genet.* **109**, 569–575 (2001).
- Yu, A. *et al.* Comparison of human genetic and sequence-based physical maps. *Nature* **409**, 951–953 (2001).
- Nobrega, M. A., Ovcharenko, I., Afzal, V. & Rubin, E. M. Scanning human gene deserts for long-range enhancers. *Science* **302**, 413 (2003).
- Glazko, G. V., Koonin, E. V., Rogozin, I. B. & Shabalina, S. A. A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions. *Trends Genet.* **19**, 119–124 (2003).
- Schwartz, S. *et al.* MultiPipMaker and supporting tools: Alignments and analysis of multiple genomic DNA sequences. *Nucleic Acids Res.* **31**, 3518–3524 (2003).
- Lim, L. P. *et al.* The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17**, 991–1008 (2003).
- Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792 (1995).
- Hattori, E. *et al.* Polymorphisms at the G72/G30 gene locus, on 13q33, are associated with bipolar disorder in two independent pedigree series. *Am. J. Hum. Genet.* **72**, 1131–1140 (2003).
- Chumakov, I. *et al.* Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia. *Proc. Natl Acad. Sci. USA* **99**, 13675–13680 (2002).
- Rowntree, C. *et al.* Deletion analysis of chromosome 13q14.3 and characterisation of an alternative splice form of LEU1 in B cell chronic lymphocytic leukemia. *Leukemia* **16**, 1267–1275 (2002).
- Armitage, J. O. & Weisenburger, D. D. New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. *J. Clin. Oncol.* **16**, 2780–2795 (1998).
- Neat, M. J. *et al.* Localisation of a novel region of recurrent amplification in follicular lymphoma to an approximately 6.8 Mb region of 13q32–33. *Genes Chromosom. Cancer* **32**, 236–243 (2001).
- Beyer, K. *et al.* Association and linkage of atopic dermatitis with chromosome 13q12–14 and 5q31–33 markers. *J. Invest. Dermatol.* **115**, 906–908 (2000).
- Kimura, K. *et al.* Linkage and association of atopic asthma to markers on chromosome 13 in the Japanese population. *Hum. Mol. Genet.* **8**, 1487–1490 (1999).
- Zhang, Y. *et al.* Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. *Nature Genet.* **34**, 181–186 (2003).
- Bentley, D. R. *et al.* The physical maps for sequencing human chromosomes 1, 6, 9, 10, 13, 20 and X. *Nature* **409**, 942–943 (2001).
- Osoegawa, K. *et al.* A bacterial artificial chromosome library for sequencing the complete human genome. *Genome Res.* **11**, 483–496 (2001).
- International Human Genome Mapping Consortium. A physical map of the human genome. *Nature* **409**, 934–941 (2001).
- Hubbard, T. *et al.* The Ensembl genome database project. *Nucleic Acids Res.* **30**, 38–41 (2002).
- Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
- Boeckmann, B. *et al.* The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**, 365–370 (2003).
- Wain, H. M., Lush, M., Ducluzeau, F. & Povey, S. Genew: the human gene nomenclature database. *Nucleic Acids Res.* **30**, 169–171 (2002).
- Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
- Stoesser, G. *et al.* The EMBL Nucleotide Sequence Database: major new developments. *Nucleic Acids Res.* **31**, 17–22 (2003).
- Schuler, G. D. Sequence mapping by electronic PCR. *Genome Res.* **7**, 541–550 (1997).
- Ning, Z., Cox, A. J. & Mullikin, J. C. SSAHA: a fast search method for large DNA databases. *Genome Res.* **11**, 1725–1729 (2001).
- Schwartz, S. *et al.* Human-mouse alignments with BLASTZ. *Genome Res.* **13**, 103–107 (2003).
- Roest Crollius, H. *et al.* Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nature Genet.* **25**, 235–238 (2000).
- Down, T. A. & Hubbard, T. J. Computational detection and location of transcription start sites in mammalian genomic DNA. *Genome Res.* **12**, 458–461 (2002).
- Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank H. M. Wain, R. C. Lovering, E. A. Bruford, M. J. Lush, V. K. Khodiyar, C. C. Talbot and S. Povey from the HUGO Gene Nomenclature Committee for official gene nomenclature, H. Reithman and R. Sudbrak for provision of clones, Y. Chen for assistance with the SNP mapping and the Wellcome Trust for financial support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.D. (ad1@sanger.ac.uk). Accession numbers for the sequences analysed for this paper can be found in Supplementary Fig. S1. All reported DNA sequences have been deposited in EMBL or GenBank.

The DNA sequence and biology of human chromosome 19

Jane Grimwood¹, Laurie A. Gordon^{2,3}, Anne Olsen^{2,3}, Astrid Terry², Jeremy Schmutz¹, Jane Lamerdin^{2,3}, Uffe Hellsten², David Goodstein², Olivier Couronne², Mary Tran-Gyamfi^{2,3}, Andrea Aerts², Michael Altherr^{2,4}, Linda Ashworth^{2,3}, Eva Bajorek¹, Stacey Black¹, Elbert Branscomb^{2,3}, Sean Caenepeel², Anthony Carrano^{2,3}, Chenier Caoile¹, Yee Man Chan¹, Mari Christensen^{2,3}, Catherine A. Cleland^{2,4}, Alex Copeland², Eileen Dalin², Paramvir Dehal², Mirian Denys¹, John C. Detter², Julio Escobar¹, Dave Flowers¹, Dea Fotopulos¹, Carmen Garcia¹, Anca M. Georgescu^{2,3}, Tijana Glavina², Maria Gomez¹, Eidelyn Gonzales¹, Matthew Groza^{2,3}, Nancy Hammon², Trevor Hawkins², Lauren Haydu¹, Isaac Ho², Wayne Huang², Sanjay Israni², Jamie Jett², Kristen Kadner², Heather Kimball², Arthur Kobayashi^{2,3}, Vladimir Larionov⁵, Sun-Hee Leem⁵, Frederick Lopez¹, Yunian Lou², Steve Lowry², Stephanie Malfatti^{2,3}, Diego Martinez², Paula McCready^{2,3}, Catherine Medina¹, Jenna Morgan², Kathryn Nelson^{2,4}, Matt Nolan², Ivan Ovcharenko^{2,3}, Sam Pitluck², Martin Pollard², Anthony P. Popkie², Paul Predki², Glenda Quan^{2,3}, Lucia Ramirez¹, Sam Rash², James Retterer¹, Alex Rodriguez¹, Stephanie Rogers¹, Asaf Salamov², Angelica Salazar¹, Xinwei She⁶, Doug Smith², Tom Slezak^{2,3}, Victor Solovyev², Nina Thayer^{2,4}, Hope Tice², Ming Tsai¹, Anna Ustaszewska², Nu Vo¹, Mark Wagner^{2,3}, Jeremy Wheeler¹, Kevin Wu¹, Gary Xie^{2,4}, Joan Yang¹, Inna Dubchak², Terrence S. Furey⁷, Pieter DeJong⁸, Mark Dickson¹, David Gordon⁹, Evan E. Eichler⁶, Len A. Pennacchio², Paul Richardson², Lisa Stubbs^{2,3}, Daniel S. Rokhsar², Richard M. Myers¹, Edward M. Rubin² & Susan M. Lucas²

¹Stanford Human Genome Center, Department of Genetics, Stanford University School of Medicine, 975 California Avenue, Palo Alto, California 94304, USA

²DOE's Joint Genome Institute, 2800 Mitchell Avenue, Walnut Creek, California 94598, USA

³Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, USA

⁴Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁵Laboratory of Biosystems and Cancer, National Cancer Institute, NIH, Bethesda, Maryland 20892, USA

⁶Department of Genetics, Center for Computational Genomics and Center for Human Genetics, Case Western Reserve University School of Medicine and University Hospitals of Cleveland, Cleveland, Ohio 44106, USA

⁷Center for Biomolecular Science and Engineering, University of California, Santa Cruz, California 95064, USA

⁸Children's Hospital Oakland, Oakland, California 94609, USA

⁹Howard Hughes Medical Institute at the Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA

Chromosome 19 has the highest gene density of all human chromosomes, more than double the genome-wide average. The large clustered gene families, corresponding high G+C content, CpG islands and density of repetitive DNA indicate a chromosome rich in biological and evolutionary significance. Here we describe 55.8 million base pairs of highly accurate finished sequence representing 99.9% of the euchromatin portion of the chromosome. Manual curation of gene loci reveals 1,461 protein-coding genes and 321 pseudogenes. Among these are genes directly implicated in mendelian disorders, including familial hypercholesterolaemia and insulin-resistant diabetes. Nearly one-quarter of these genes belong to tandemly arranged families, encompassing more than 25% of the chromosome. Comparative analyses show a fascinating picture of conservation and divergence, revealing large blocks of gene orthology with rodents, scattered regions with more recent gene family expansions and deletions, and segments of coding and non-coding conservation with the distant fish species *Takifugu*.

The finished human chromosome 19 sequence, comprising a gene density more than double the genome-wide average^{1,2}, marks the culmination of 18 yr of research spanning the history of modern genomics. The Department of Energy's (DOE) role in the project was initiated through an effort to understand how the body responds to and repairs radiation damage^{3,4}. A link between unrepaired DNA damage and human cancer⁵, and the subsequent mapping of multiple DNA repair genes to the chromosome⁶, provided the impetus for the DOE to select the chromosome as one of its sequencing targets⁷. The physical mapping was completed at Lawrence Livermore National Laboratory (LLNL) in 1995 with the generation of the first high-quality metric map of a human chromosome⁸. In 1999, sequencing and finishing was transferred from LLNL to the DOE's Joint Genome Institute's Production Sequencing Facility and the Stanford Human Genome Center.

The clone map and finished sequence

A complete *EcoRI* restriction map of chromosome-19-specific cosmids⁹ provided the foundation for sequencing human chromosome 19. A physical metric across the entire map was generated by estimating the distances between specific points in the cosmid contigs, using fluorescence *in situ* hybridization in sperm pro-

nuclei^{10,11}. These initial efforts provided a metrically resolved scaffold framework of 216 cosmid reference points defining contig order and orientation, gap locations and gap sizes, and efficiently directed future map closure efforts (see Supplementary Methods). Gaps in this map were subsequently filled by screening 100-fold clone coverage of the genome using available large-insert genomic libraries (see Supplementary S1). The final tiling path consists of 860 clones that span the entire euchromatin regions of both the p and q arms of the chromosome. The path consists of 511 chromosome-19-specific cosmids, 333 bacterial artificial chromosomes, 11 fosmids, three P1-derived artificial chromosomes, one yeast artificial chromosome clone and one genomic polymerase chain reaction (PCR) product. The chromosome is represented in four contigs, one of which covers the entire q arm, and the clone map is estimated to cover 99.9% of the euchromatin sequence.

Sequence was generated using a clone-by-clone shotgun sequencing strategy¹ followed by finishing using a custom primer approach. Recalcitrant areas or hard gaps were closed with additional sequence data derived from transposon sequencing, small insert shatter libraries or PCR. Each clone was finished according to the agreed international standard for the human genome (<http://genome.wustl.edu/Overview/g16stand.php>). On the basis of internal and

external quality checks, we estimate the accuracy of our finished sequence to exceed 99.99%¹². In total, we finished 55,785,651 base pairs (bp) and estimate the total size of the chromosome, including the two clone gaps and the recalcitrant centromeric and subtelomeric regions, to be 63.8 megabases (Mb).

Comparison to physical and genetic maps

We observed strong concordance between the chromosome 19 sequence and previously existing physical and genetic maps. All sequence-tagged sites from the Genethon microsatellite-based genetic map¹³, the deCODE map¹⁴ and the Marshfield genetic maps¹⁵ were present in the chromosome 19 sequence (see Supplementary Methods).

We compared recombination distances in the deCODE female, male and sex-averaged meiotic maps¹⁴ with physical distance as determined from the sequence assembly. Recombination statistics for chromosome 19 are similar to other human chromosomes, with the female and sex-averaged comparisons showing a relatively linear relationship between recombination and physical distances, with an average of 2.1 cM Mb⁻¹ (Fig. 1). The male meiotic map, however, shows striking differences, particularly in the q arm, showing a long 'desert' with a meiotic distance of only 2.7 cM in the 20.7–44.1 Mb euchromatic region surrounding the centromere (<0.18 cM Mb⁻¹). Although the basis for this result is not clear, it is interesting that this segment of the chromosome is particularly rich in long interspersed nuclear element (LINE) sequences, a finding in agreement with other chromosomes¹⁶. Also consistent with other chromosomes is a marked increase in male recombination near both telomeres¹⁶.

The chromosome landscape

A hallmark feature of chromosome 19 is its unusually high density of genes. On average, 26 protein-coding gene loci were found per megabase, and exons cover 3.55 Mb of the sequence (6.4%), a percentage significantly higher than the genome-wide average of 1.5%². Protein-coding loci (exons plus introns) span 28.1 Mb (50% of the chromosome) and the average annotated mature transcript contains 58% and 42% coding and untranslated regions, respectively.

Chromosome 19 is also unusual in its density of repeat sequences. Nearly 55% of this chromosome consists of repetitive elements (Table 1), whereas chromosomes 6, 7, 14, 20, 21 and 22 all have repeat contents ranging from 40% to 46% (the genome average is 44.8%)^{1,17–22}. This difference is due mainly to an unusually high content of short interspersed nuclear elements (SINEs) on chromo-

some 19. Specifically, *Alu* repeats make up 25.8% of the chromosome, compared with 13.8%, 13.3%, 9.5% and 16.8% on chromosomes 7, 14, 21 and 22, respectively.

In addition, the G+C content of the chromosome is unusually high, with an average of 48%. This compares with 41% as reported in the whole human genome analysis¹. G+C content can be separated into two distinct chromosomal categories: regions of human-specific gene family expansions and regions of 1:1 gene orthology blocks with mouse (Fig. 2). In the 20 large duplicated gene family regions (covering a combined 15 Mb of the chromosome) the average G+C content is 43.1%. In contrast, the G+C content is significantly higher (50.3%) in the remaining 38 Mb of the chromosome where there is a clear 1:1 human/mouse orthologous relationship (see 'Comparative biology' section). This high G+C content correlates with gene and *Alu* repeat density, and is negatively correlated with LINE density. Finally, two-thirds of genes have at least one CpG island and 1,623 (88%) of these 1,841 elements are found within 2,000 bp upstream to 1,000 bp downstream of the putative transcription start sites of annotated genes.

The gene and protein catalogue of chromosome 19

An automated pipeline of evidence-based and *ab initio* methods was used to place gene model transcripts on the underlying genomic sequence. Subsequent to this, each transcript was manually reviewed using a combination of human-expressed sequence evidence (messenger RNA and expressed sequence tags (ESTs)) and homology to known genes in humans, mice and other organisms. Additional genes were identified manually from underlying experimental data. Ultimately, a total of 1,461 protein-coding regions were verified as valid gene loci (see Supplementary S2). These loci contain 2,341 full-length (or nearly full-length) transcripts, as well as partial evidence for additional splice variants as discussed below. We placed loci in the following three categories: (1) 'known' genes that correspond to RefSeq genes²³, human complementary DNA or protein sequences; (2) 'novel' or previously unidentified loci that have an open reading frame (ORF) greater than 100 amino acids, and/or are identical to a spliced human EST, and/or have homology to known genes or proteins (all species); and (3) 'pseudogenes', which have sequence similar to genes or proteins found elsewhere in the genome but lack the introns present in the original version (processed) and/or have a disrupted or partial ORF. Transcripts for which a unique ORF could not be determined and putative genes (*ab initio* models) with no supporting experimental evidence were not considered valid.

We identified 1,320 known loci based on 1,551 RefSeq genes and other nearly full-length cDNA sequences in GenBank that mapped to chromosome 19. On the basis of EST evidence, we were able to extend 60% of these RefSeq transcripts by more than 50 nucleotides at the 5' and/or 3' ends while maintaining the original ORF. A total of 41% of the RefSeq loci were extended at the 5' end, more accurately locating the transcriptional start site for these transcripts.

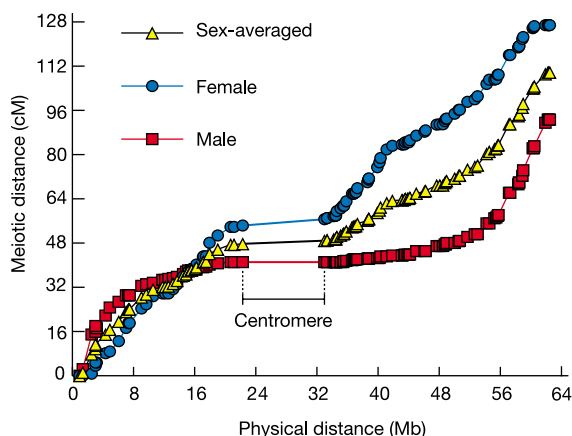


Figure 1 Chromosome 19 meiotic distance versus sequence-based physical distance. The genetic and physical maps were aligned from the telomere of the short arm to the telomere of the long arm. The position of each genetic marker on the female, male and the sex-averaged genetic map is indicated.

Table 1 Interspersed repetitive elements

Element	Coverage (bp)	Coverage (%)
Alu	14,415,071	25.83
L1	5,551,771	9.95
L2	1,215,945	2.18
LTR	1,178,970	2.11
MEI	1,837,372	3.29
MIR	864,988	1.55
HERV	1,057,852	1.90
MLT	9,270,34	1.66
Simple	781,731	1.4
Low complexity	439,717	0.79
Other	2,838,415	5.09
Total	31,108,866	55.75

HERV, human endogenous retrovirus-like; LTR, long terminal repeat; MEI, medium reiterated frequency repeat; MIR, mammalian interspersed repeat; MLT, mammalian LTR transposon.

A total of 88% of the transcripts end with a stop in the final exon/untranslated region.

We found evidence for 141 novel loci for which RefSeq genes were not available. These loci are supported by other nearly full-length cDNA sequences, one or more spliced ESTs and/or similarity to known human or mouse sequences. Within this group are 58 human loci modelled using orthologous mouse cDNA sequences. Transfer RNA genes are one of the best-understood non-protein-coding RNA genes with respect to function. We confidently predicted 11 tRNA genes and three tRNA pseudogenes, in stark contrast with the 157 tRNAs found on the p arm of chromosome 6 (ref. 17).

The largest gene on chromosome 19 is the alpha 1A subunit of the P/Q-type voltage-dependent calcium channel (*CACNA1A*), which extends over more than 300 kilobases (kb) and contains 47 exons. The transcript with the most exons is the skeletal muscle ryanodine

receptor (*RYR1*), which has 105 exons spread over nearly 154 kb. The largest exon on the chromosome is 21,693 bp and is found within the *MUC16* gene, a gene encoding an unusually large transmembrane glycoprotein with a role in embryonic development and neoplastic transformation²⁴.

Alternative splicing

We characterized the extent of alternative splicing based on the existing cDNA/EST data. Considering only mRNA sequences in GenBank, we identified 2,341 distinct chromosome 19 transcripts that provided an average coverage of 1.6 annotated transcripts per locus (see Supplementary S2). These mRNAs provide strong evidence for alternative splicing of 568 (39%) of the 1,461 annotated gene loci, with each having two or more associated transcripts. Furthermore, an additional 452 genes have at least one EST sequence overlapping with non-annotated exons and also contain

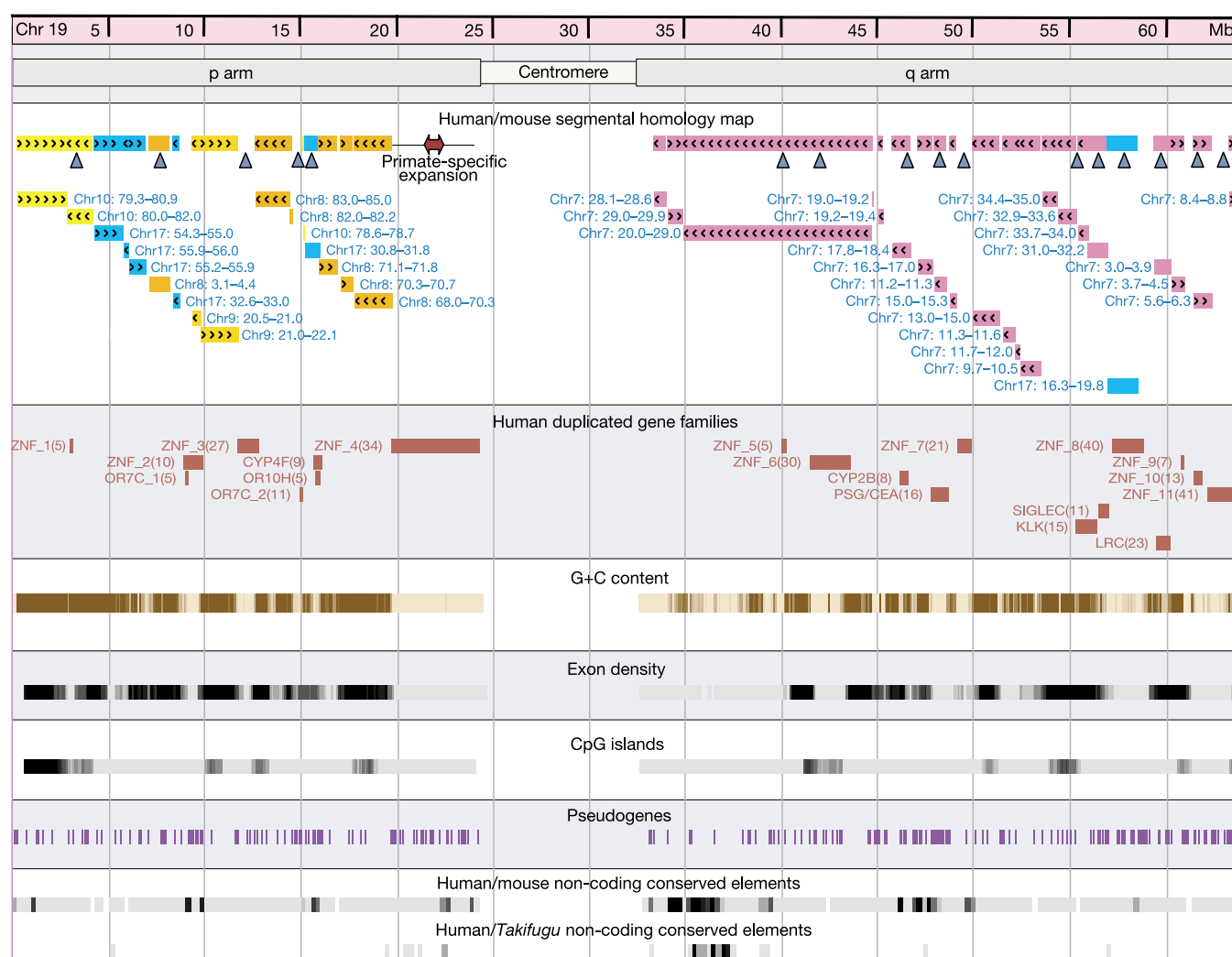


Figure 2 The chromosome 19 landscape. The following sections appear in order from top to bottom: (1) human/mouse blocks of homology larger than 100 kb. The grey triangles indicate regions where significant lineage-specific and no 1:1 orthology can be identified. (2) Same as (1) but showing positions in the mouse genome. This map shows extensive intrachromosomal and interchromosomal rearrangement between human and mouse. (3) The duplicated gene families covering more than 25% of the chromosome (see 'Gene families and duplication analysis' section). Numbers in parentheses are the number of functional genes in the cluster (see also Table 2). (4) G+C content (see 'Chromosome landscape' section). Note that most of the zinc finger and olfactory receptor gene families have a lower G+C content whereas the rest of the chromosome has an unusually high

G+C content (although the G+C content is not high in the corresponding mouse syntenic regions). (5) Exon density using a 1-Mb sliding window. The exon density correlates with the G+C content fluctuations. (6) The location of CpG islands on chromosome 19 (see 'Chromosome landscape' section). (7) Pseudogenes identified during the annotation of chromosome 19 (see 'Pseudogene' section). (8) Human/mouse and human/*Takifugu* non-coding DNA element density. This density is uneven over the chromosome with the 5-Mb region in the proximal portion of the q arm containing the highest density of human/mouse non-coding elements and the majority of the non-coding human/*Takifugu* elements. These regions have a low G+C content, a low gene density and fewer mouse/human breakpoints (see 'Comparative biology' section).

flanking canonical splice sites at the genomic locus. Thus, existing experimental data support alternative splicing for a minimum of 1,020 of the genes (70%) on chromosome 19.

It is likely that an even larger fraction of chromosome 19 genes are subject to some form of alternative splicing. As most of our conclusions are based on existing transcribed sequence data, the depth of the EST database seems to be a limiting factor. In fact, of the 184 genes with a total of 500 or more overlapping ESTs, 181 (98%) displayed low-frequency alternative transcripts. Thus deeper EST clone coverage would probably show that a very large fraction of loci can exhibit alternate splicing. Recent studies support this conclusion²⁵.

Pseudogenes

We identified 321 pseudogenes on chromosome 19 (see Supplementary S3). Of these, 177 (55%) were classified as 'processed' pseudogenes, that is, products of viral retrotransposition events involving spliced messenger RNAs that can frequently be identified by the absence of introns that are present in the parent locus and by the presence of poly(A) stretches embedded in the adjacent genomic sequence. Forty-six (14%) pseudogenes probably arose from genomic duplication events, displaying remnants of introns from the parent locus. An additional 98 (31%) pseudogenes were classified as

potential pseudogene fragments owing to their partial nature.

A total of 70 (22%) of the 321 pseudogenes on chromosome 19 contain uninterrupted, but partial, ORFs and probably represent young processed pseudogenes that have not had sufficient time to accumulate mutations to disrupt their ORF, but may also have retained some function. Of the 22 olfactory receptor loci annotated as pseudogenes, four contain a complete ORF. Recent studies have shown that a significant fraction of putative olfactory receptor pseudogenes in the genome are segregating as alleles with intact, presumably functional copies in the human population²⁶. Whether any olfactory receptor, or other, pseudogenes on chromosome 19 also vary in humans between such potential functional and non-functional states remains to be explored experimentally.

Gene families and duplication analysis

Chromosome 19 is notable for the prevalence of duplication structures of two types: tandemly clustered gene families and large segmental duplications. As to the latter, chromosome 19 shows evidence of extensive genomic duplication with 7.35% of the sequence sharing sequence homology to more than one location in the genome (Fig. 3; see also Supplementary S4). In contrast, whole-genome estimates of segmental duplication predict 5–6% duplicated sequence using the same alignment parameters (≥ 1 kb length; $\geq 90\%$ sequence identity)¹. This enrichment on chromosome 19 is predominantly due to an increase in intrachromosomal duplications (6.20% of the sequence) rather than interchromosomal duplications (1.69% of the sequence). Using sequence divergence as a surrogate of evolutionary age, these data indicate that most of the tandem expansions of duplications on chromosome 19 occurred much earlier (30–40 million years ago) when compared with the more recent interchromosomal duplication events. The most marked feature of the segmental duplication pattern on chromosome 19 is the pattern of large intrachromosomal duplications ($>90\%$ sequence identity) clustered in tandem arrangement (Fig. 3; see also Supplementary S5).

More than 25% of the genes on chromosome 19 are members of large, well-defined, tandemly clustered gene families (Fig. 4 and Table 2). Considerable evidence has documented the existence of lineage-specific changes within these and other tandemly clustered families due to ongoing gene duplication, deletion and mutational

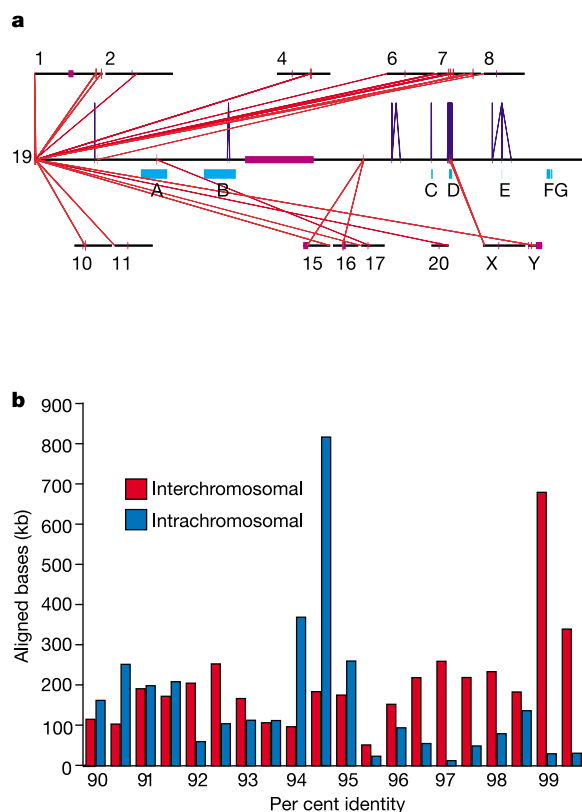


Figure 3 Recent segmental duplications on chromosome 19. **a**, Large (>20 kb), highly similar ($>95\%$) intrachromosomal (blue) and interchromosomal (red) segmental duplications are shown for chromosome 19. Chromosome 19 is drawn at a greater scale relative to the other chromosomes. Gene clusters detected in duplicated sequences (>1 kb with identity $>90\%$) are represented as light blue bars below the chromosome 19 sequence. A, B, ZNF genes; C, cytochrome P450; D, pregnancy-specific α -1-glycoprotein (PSG); E, chorionic gonadotropin β -peptide (CGB); F, leukocyte immunoglobulin-like receptor; G, killer cell immunoglobulin-like receptor. **b**, Sequence similarity of segmental duplications. For all pairwise alignments, the total number of aligned bases was calculated and binned based on per cent sequence identity. Sequence identity distributions for interchromosomally (red) and intrachromosomally (blue) duplicated bases are shown.

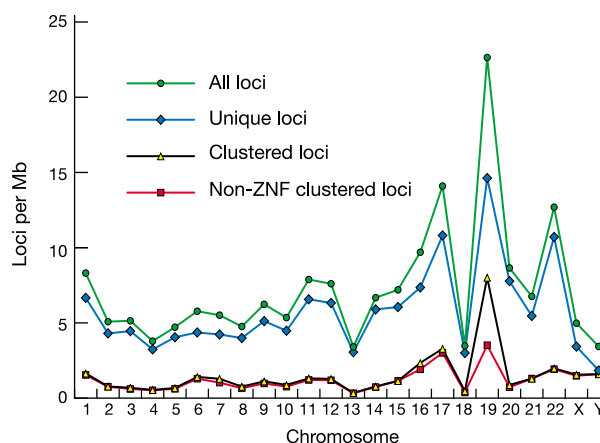


Figure 4 Chromosome 19 is extremely gene dense relative to other human chromosomes. Known genes (November 2003) were downloaded from the UCSC browser (<http://genome.ucsc.edu>) and plotted relative to the relative chromosome size (minus the centromere). Yellow triangles (clustered loci) indicate genes in tandem duplications, whereas blue diamonds (unique loci) indicate loci that are not duplicated in a tandem manner. Green circles (all loci) represent the sum of clustered and unique loci. A subset of the clustered loci that does not involve the KRAB-Kruppel ZNF genes is shown as red squares (non-ZNF clustered loci). As well as having the highest number of genes, chromosome 19 also has the largest number of genes contained in tandem gene families.

events^{27–34}. These clustered sets of paralogues therefore represent a potentially rich source of genetic diversity, and because of their prevalence, chromosome 19 has an especially dynamic evolutionary history. The largest group of such genes on chromosome 19 encodes Krüppel-type (or C2H2) zinc finger transcription factor (ZNF) proteins, with 266 of the approximately 800 total human genes of this type located primarily within 11 large familial clusters^{27,29}. Chromosome 19 contains members of several different ZNF gene subfamilies but most of the clustered genes belong to the KRAB-ZNF subtype (Table 2). One large cluster of ZNF genes, located in the pericentromeric region of the short arm, is exclusive to and arose early in the primate lineage³⁰. A unique aspect of this region is the admixture of classical centromeric β -satellite sequences with the ZNF genes. A total of 27 blocks of β -satellite repeat (each ranging from 10 to 50 kb; mean = 22 ± 8 kb) map immediately distal to the centromeric α -satellite sequence. These blocks are located throughout the first 4 Mb of the pericentromeric region of 19p12 with an average of 114 kb separating each β -satellite block. Embedded between most of these β -satellite blocks are 1 to 2 KRAB-ZNF genes, indicating that these structures were coordinately duplicated³⁰.

Human chromosome 19 also carries a large collection of genes encoding receptor proteins with immunoglobulin-like domains. Genes of the closely related leukocyte immunoglobulin-like receptors (LILRA and LILRB), the leukocyte-associated immunoglobulin-like receptors (LAIR) and the killer cell immunoglobulin receptors (KIR) are found together in the leukocyte cluster region (LCR) of 19q13.4. The proteins encoded by these loci function as receptors for specific classes of antigens on the surface of various types of immune cells. As with the ZNF and olfactory receptor (OR) genes, the LAIR, LILR and KIR gene families differ extensively in their relative numbers and types between different vertebrate lineages. The variety of different immunoglobulin-like receptor proteins may define some of the major differences in strategies adopted by particular lineages to combat infectious agents and antigens encountered in different environments³¹. The KIR gene family arose recently in the primate lineage, and consistent with this, the repertoire of this gene family varies both in gene number and type even between individual humans. Specific KIR haplotypes have been shown to determine differential susceptibility to immune-related diseases, and are associated with differential rates of pro-

gression to AIDS in HIV-infected individuals³². The KIR haplotype represented in the public human genome sequence corresponds to the most commonly occurring haplotype in the Caucasian population, called A-1D (ref. 32). This carries nine of the seventeen previously described KIR family members, including a known deletion variant of the *KIRDS4* locus, *KIR2DS4*.

Several other large and evolutionarily diverse gene clusters exist on the chromosome, all with diverse evolutionary histories and involvement in various medical conditions. In addition to the LRC family genes, of particular medical importance are the rapidly evolving cytochrome P450 subfamily II genes (CYP2)³³ involved in the metabolism of steroid hormones, carcinogens and other substances, and the kallikrein (KLK)³⁴ serine protease family, associated with tumour progression. The positions, size and functions of these gene families are summarized in Table 2.

Comparative biology

To define further the chromosomal landscape and annotate putative functional sequence elements, we performed a comparative analysis of finished human chromosome 19 versus the draft mouse and *Takifugu* (fish) genomes. Using DNA alignments we constructed a homology map refining the map of syntenic homology between chromosome 19 and the mouse genome. Alignments of chromosome 19 with orthologous mouse sequence reveals regions of two general types: vast regions of synteny with 1:1 gene orthology, and significant segments where, owing to lineage-specific duplication and rearrangement events, 1:1 orthology is rare. Intervals in which unambiguous 1:1 orthologous relationships can be defined span over 45 Mb or 82% of chromosome 19 (excluding the centromere). By scanning these regions for contiguous collinear nucleotide similarity, 38 blocks larger than 100 kb were identified, the longest segment being 9.7 Mb (Fig. 2; see also Supplementary S6). All of these blocks lie within the boundaries of larger, previously characterized segments of human–mouse syntenic homology²⁷. Homology scanning and manual curation permitted a more precise definition of syntenic borders and identified previously undetected internal rearrangements within the larger segments. Within the p arm, extensive rearrangements have occurred since the time of primate–rodent divergence, with 16 major rearrangements located in nine distinct syntenic segments derived from four mouse chromosomes (mouse chromosomes 8, 9, 10 and 17). In contrast,

Table 2 Chromosome 19 gene family clusters

Family/subfamily	Number of functional genes	Chromosome 19 location (Mb)	Function
Krüppel-type zinc finger genes	266		Transcriptional activation and repression
KRAB-ZNF subfamily	202		
Other subfamilies	64		
Cluster 1	5	2.7–2.9	
Cluster 2	10	8.7–9.7	
Cluster 3	27	11.5–12.6	
Cluster 4	34	19.5–24.1	
Cluster 5	5	34.8–35.1	
Cluster 6	30	36.3–38.4	
Cluster 7	21	44.0–44.7	
Cluster 8	40	52.0–53.7	
Cluster 9*	7	55.6–55.8	
Cluster 10	13	56.2–56.7	
Cluster 11	41	57.0–58.7	
Olfactory receptors			Odorant detection
Cluster 1	5	8.8–9.0	
Cluster 2	11	14.7–14.9	
OR10H family	5	15.6–15.8	
Cytochrome P450			Leukotriene receptors/inflammation response
CYP4F family	9	15.5–15.9	
CYP2 family	8	40.9–41.4	Metabolism of drugs, toxins and steroid hormones
Pregnancy-specific glycoproteins/carcinoembryonic antigens	16	42.6–43.5	
Sialic acid glycoprotein lectins	11	51.3–51.8	Maintenance of pregnancy
Kallikrein proteins	15	50.1–51.2	Sialic-acid-recognizing cell surface lectins
Immunoglobulin-like receptors LAIR, LILR, ILT and KIR	23	54.3–55	Tissue-specific serine proteases
			HLA antigen receptors on T- and B-cell surfaces

* All ZNF clusters except cluster 9 contain at least some KRAB-ZNF family members and many clusters contain ZNF and mixed types; cluster 9 is comprised entirely of SCAN-ZNF genes.

the q arm aligns almost entirely with mouse chromosome 7 and rat chromosome 1, the single exception being a 3-Mb region related to mouse chromosome 17. However, within this relatively stable synteny group are found more than 21 intrachromosomal breakpoints involving transposition or inversion events that have occurred since the primate–rodent divergence (Fig. 2).

The second type of chromosome 19 regions corresponds to tandem gene families, in which extensive lineage-specific gene duplication and loss have occurred (Table 2). Reflecting the recent duplication and high divergence rates within these clusters, only approximately 51% of gene-family exons display an apparent match in the mouse or the rat genomes compared with about 84.7% when the human genome is interrogated for matches³⁵. In some of these rapidly evolving regions, especially the ZNF and OR clusters from which both dispersed and tandem gene copies have been generated independently in each lineage, human–mouse homology relationships are less obvious. This picture is complicated by the fact that many chromosome 19 gene families are localized at the boundaries of human–rodent syntenic rearrangements; in several cases, family members have been split by these events onto separate chromosomes in mouse²⁷. However, members of most conserved chromosome 19 gene clusters are clearly anchored within syntenically conserved positions by identities of flanking unique genes and evolutionary analyses^{27–29}.

Comparisons of human chromosome 19 and mouse DNA sequences confirm extensive coding and non-coding conservation with 4,586 discrete fragments showing conservation (>70% identity with a score of match-mismatch >60) but lacking evidence of encoding protein or being transcribed. The putative conserved non-coding elements are unevenly distributed along the chromosome with several high-density clusters in the proximal portion of the q arm in a large region with low G+C content, low gene density and low mouse/human breakpoint density (Fig. 2). It is unclear whether the majority of these are functional elements under strong evolutionary pressure or whether they are simply conserved owing to a low local mutation rate.

Additional DNA comparisons of human chromosome 19 against *Takifugu* revealed extensive coding and significantly less non-coding conservation. A total of 57% of human exons are covered by *Takifugu* conservation. In contrast, only 66 chromosome 19 sequences were conserved with *Takifugu* (>70% identity and match-mismatch score >50) for which no coding or transcribed evidence could be found. Three sequences showed evidence of being non-coding RNA genes³⁶ (<http://www.genetics.wustl.edu/eddy/software/#qrna>) whereas the remaining 63 appear to be untranscribed in nature. Similar to human–mouse non-coding conserved elements, these human–*Takifugu* non-coding conserved elements are also enriched in the low-gene-dense proximal portion of the q arm. This finding of ancient evolutionary constraints on a small fraction of non-coding chromosome 19 DNA suggests a probable role in basic vertebrate biological functions, and recent work has shown that a significant fraction of non-coding elements conserved between human and *Takifugu* can have gene regulatory activity even when located great distances from genes³⁷.

Conclusions and implications for human disease

From the beginning of the Human Genome Project (HGP), one of the major goals was to facilitate the identification and study of the functions of genes that are involved in genetic diseases, both simple mendelian forms and those with more complex inheritance. Now that the sequence is complete, the process of mapping and identifying a disease gene is no longer limited by experimental molecular biology but essentially only by the number of meioses in accurately diagnosed families. As chromosome 19 is crowded with genes, it is not surprising that it has a large number of well-mapped and cloned genes for single-gene disorders, as well as many loci with evidence for association with complex traits. Currently, there are at least 97

single-gene mendelian traits, most of which correspond to rare genetic diseases, that have been localized to specific loci or regions on the chromosome by meiotic mapping in families (see Supplementary S7 and <http://www.ncbi.nlm.nih.gov/Omim/>).

The genes and associated mutations have been identified for about 75% of these traits. Some of these genes were identified through classical biochemical and molecular approaches before the formal start of the HGP, including the low-density lipoprotein receptor gene³⁸, which leads to familial hypercholesterolaemia when mutated³⁹, and the gene encoding the erythropoietin receptor, which is important for erythroid and myeloid cell differentiation and results in autosomal-dominant erythrocytosis when mutations that increase its expression are present⁴⁰. However, most of the genes contributing to altered phenotypes were found with positional cloning approaches that became increasingly facile as the maps and sequence of the chromosome were produced and refined. Even so, there remain at least 20 mendelian diseases mapped on chromosome 19 for which the genes have not yet been identified. Finally, the neurturin gene lying on chromosome 19, which encodes the ligand of a tyrosine kinase receptor (RET), represents one of the limited examples of a gene that has convincingly been demonstrated to contribute to a multigenic trait. Mutations in both the neurturin and the receptor genes together lead to Hirschsprung's disease, whereas mutations in neurturin alone are insufficient to cause the disorder⁴¹. Biology has truly entered a new era with the completion of the sequence of the entire human genome. The finished sequence of chromosome 19 will facilitate the identification of additional genes contributing to single-gene disorders as well as complex traits. In addition, the large number of evolutionarily conserved non-coding sequences shared with other vertebrates suggest new candidate regions for searching for the genetic basis of human disorders and quantitative traits where sequence alterations of gene regulatory elements have been suggested as a frequent molecular mechanism⁴². □

Methods

Annotation

All human and mouse expressed sequences in GenBank (February 2003) were aligned to their 'best-in-genome position' against repeat-masked chromosome 19 sequence (version 1, February 2003) using BLAT⁴³. Alignments were post-processed to correct alignment errors and enforce common splice signals⁴⁴. The mRNA and EST evidence was analysed to retain information about original noted start, CDS, protein sequence and any indicated poly(A) site/signal. In addition, putative transcripts were produced using the evidence-based gene finder FgenesH++ (ref. 44) and by aligning syntenic mouse cDNAs to the chromosome with GeneWise⁴⁵. cDNA and computational predictions that consistently overlapped ESTs at the same locus were automatically extended. Other analyses including MZFEF, GenomeScan and CpGseek were run in parallel. Cases where a human mRNA was wholly contained within another human mRNA as a distinct transcript were not reported. All results were loaded into a MySQL database. Finally, each predicted model was analysed for domain content with InterPro and aligned using a double affine Smith–Waterman algorithm against human and mouse proteomes, as well as Swissprot and other GenBank proteins. The resulting gene structures were inspected relative to EST evidence and protein homology/domain content using a web-based interface. With a combination of the browser, local tools and the Apollo editing system⁴⁶, a distributed group of annotators corrected apparent errors as necessary. For gene families with well-defined domain content, custom gene models were constructed with respect to the known structures of these gene families, using the following hierarchy of criteria: (1) matches to cDNA sequence data; (2) protein homology; and (3) gene prediction data. If there was evidence for additional genes or transcripts that were not previously represented by the auto-promoted models, they were also indicated. We report only those loci with model transcripts confirmed by two or more methods. After the completion of the manual curation, the resultant gene catalogue was aligned to the NCBI July 2003 build 34 of chromosome 19. The browser interface can be found at <http://genome.jgi-psf.org/Chr19/Chr19.home.html>.

Pseudogenes were identified by means of a two-pronged approach. First, the gene structure (exon number) of all FgenesH models was compared with the gene structure of the models' closest human homologues. Those FgenesH models containing fewer exons than their closest human homologue were deemed potential processed or partial pseudogenes. Each of these sequences was manually analysed for pseudogene status. The chromosome 19 sequence was then masked of all repeats (<http://ftp.genome.washington.edu/RM/RepeatMasker.html>) and of all coding exons from the 1,461 protein-coding genes. The most current version of the human IPI protein data set was then aligned against the masked sequence using two alignment programs: tblastn and

prot_map. Protein alignments were filtered to select the longest and best-scoring alignment for a given genomic locus with the top alignment at each locus being manually analysed to determine pseudogene status. Single-exon pseudogenes (including those interrupted by repetitive elements or short inserts, that is <100 bp) that shared similarity to a protein sequence from the most current version of the Ensembl database, such that the pseudogene sequence covers at least 70% of the coding sequence, where classified as processed pseudogenes; any single-exon pseudogenes aligning to less than 70% of the length of the closest Ensembl protein sequence were classified as potential pseudogene fragments.

Comparative analysis

The mouse and rat genomes were taken from <http://genome.ucsc.edu>; we used the repeat masked version of mouse February 2003 freeze (mm3) and the rat June 2003 freeze (rn3). The *Takifugu* genome assembly is available from the JGI website (<http://www.jgi.doe.gov>). The cross-species alignments were performed using BLASTZ⁴⁷. The rodent genomes were chopped into 1-Mb pieces with a 10-kb overlap. Each piece was then aligned to the entire chromosome 19 sequence. The raw DNA BLASTZ alignments were then used to create the segmental maps, include duplications, and extract the conserved regions using the PARAGON software (O.C., unpublished data). The mRNA and EST alignments (<http://genome.ucsc.edu>) were used to filter out the non-coding set from coding evidence (human/mouse EST, spliced and non-spliced, non-human/mouse mRNA and EST aligned to each genome using BLAT). Some of the coverage statistics data were done with software developed by J. Kent (<http://www.cse.ucsc.edu/~kent>). Parts of Fig. 2 were made using a local modified installation of the UCSC genome browser.

Segmental duplication analysis

We used a BLAST-based detection scheme⁴⁸ to identify all pairwise similarities representing duplicated regions (≥ 1 kb and $\geq 90\%$ identity) within chromosome 19 (version 1, February 2003) and compared them to all other chromosomes in the NCBI genome assembly (build 31). We compared these BLAST-based pairwise similarities to the whole-genome shotgun sequence detection database of segmental duplications, which was ascertained via an assembly-free method⁴⁸. The program Parasight (J. A. Bailey, unpublished data) was used to generate images of pairwise alignments. We also analysed pairwise alignments for per cent identity and the number of aligned bases. β -Satellite repeats were detected using RepeatMasker (15 May 2002 version) on sensitive settings (A. Smit and P. Green, unpublished data).

Received 1 December 2003; accepted 10 February 2004; doi:10.1038/nature02399.

- International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
- Setlow, R. B. & Setlow, J. K. Evidence that ultraviolet light induced thymine dimers in DNA cause biological damage. *Proc. Natl Acad. Sci. USA* **48**, 1250–1257 (1962).
- Setlow, R. B. & Carrier, W. L. The disappearance of the thymine dimers from DNA: An error correcting mechanism. *Proc. Natl Acad. Sci. USA* **51**, 226–231 (1964).
- Cleaver, J. E. Defective repair replication of DNA in xeroderma pigmentosum. *Nature* **218**, 652–656 (1968).
- Mohrenweiser, H. W. *et al.* Refined mapping of the three DNA repair genes, ERCC1, ERCC2, and XRCC1, on human chromosome 19. *Cytogenet. Cell Genet.* **52**, 11–14 (1989).
- Patrinos, A. & Drell, D. W. The Human Genome Project: view from the Department of Energy. *J. Am. Med. Womens Assoc.* **52**, 8–10 (1997).
- Ashworth, L. K. *et al.* An integrated metric physical map of human chromosome 19. *Nature Genet.* **11**, 422–427 (1995).
- de Jong, P. *et al.* Human chromosome-specific partial digest libraries in lambda and cosmid vectors. *Cytogenet. Cell Genet.* **51**, 985 (1989).
- Brandiff, B. F. *et al.* Human chromosome 19p: A fluorescence *in situ* hybridization map with genomic estimates for 79 intervals spanning 20 Mbp. *Genomics* **23**, 582–591 (1994).
- Gordon, L. A. *et al.* A 30-Mb metric fluorescence *in situ* hybridization map of human chromosome 19q. *Genomics* **30**, 187–192 (1995).
- Felsenfeld, A., Peterson, J., Schloss, J. & Guyer, M. Assessing the quality of the DNA sequence from the Human Genome Project. *Genome Res.* **9**, 1–4 (1999).
- Dib, C. *et al.* A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**, 152–154 (1996).
- Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nature Genet.* **31**, 241–247 (2002).
- Broman, K. W., Murray, J. C., Sheffield, V. C., White, R. L. & Weber, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63**, 861–869 (1998).
- Yu, A. *et al.* Comparison of human genetic and sequence-based physical maps. *Nature* **409**, 951–953 (2001).
- Mungall, A. J. *et al.* The DNA sequence and analysis of human chromosome 6. *Nature* **425**, 805–811 (2003).
- Hillier, L. W. *et al.* The DNA sequence of human chromosome 7. *Nature* **424**, 157–164 (2003).
- Hellig, R. *et al.* The DNA sequence and analysis of human chromosome 14. *Nature* **421**, 601–607 (2003).
- Deloukas, P. *et al.* The DNA sequence and comparative analysis of human chromosome 20. *Nature* **414**, 865–871 (2001).
- Hattori, M. *et al.* The DNA sequence of human chromosome 21. *Nature* **405**, 311–319 (2000).
- Dunham, I. *et al.* The DNA sequence of human chromosome 22. *Nature* **402**, 489–495 (1999).
- Pruitt, K. D. & Maglott, D. R. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.* **29**, 137–140 (2001).
- O'Brien, T. J., Beard, J. B., Underwood, L. J. & Shigemasa, K. The CA 125 gene: a newly discovered extension of the glycosylated N-terminal domain doubles the size of this extracellular superstructure. *Tumour Biol.* **23**, 154–169 (2002).
- Kan, Z., States, D. & Gish, W. Selecting for functional alternative splices in ESTs. *Genome Res.* **12**, 1837–1845 (2002).
- Menashe, I., Man, O., Lancet, D. & Gilad, Y. Different noses for different people. *Nature Genet.* **34**, 143–144 (2003).
- Dehal, P. *et al.* Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. *Science* **293**, 104–111 (2001).
- Newman, T. & Trask, B. J. Complex evolution of 7E olfactory receptor genes in segmental duplications. *Genome Res.* **13**, 781–793 (2003).
- Shannon, M., Hamilton, A. T., Gordon, L., Branscomb, E. & Stubbs, L. Differential expansion of zinc-finger transcription factor loci in homologous human and mouse gene clusters. *Genome Res.* **13**, 1097–1110 (2003).
- Eichler, E. E. *et al.* Complex β -satellite repeat structures and the expansion of the zinc finger gene cluster in 19p12. *Genome Res.* **8**, 791–808 (1998).
- Trowsdale, J. *et al.* The genomic context of natural killer receptor extended gene families. *Immunol. Rev.* **181**, 20–38 (2001).
- Hsu, K. C. *et al.* The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. *Immunol. Rev.* **190**, 40–52 (2002).
- Hoffmann, S. M., Nelson, D. R. & Keeney, D. S. Organization, structure and evolution of the CYP2 gene cluster on human chromosome 19. *Pharmacogenetics* **11**, 687–698 (2001).
- Yousef, G. M. & Diamandis, E. P. The new human tissue kallikrein gene family: structure, function, and association to disease. *Endocr. Rev.* **22**, 184–204 (2001).
- Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
- Rivas, E. & Eddy, S. R. Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics* **2**, 8 (2001).
- Nobrega, M. A., Ovcharenko, I., Afzal, V. & Rubin, E. M. Scanning human gene deserts for long-range enhancers. *Science* **302**, 413 (2003).
- Yamamoto, T. *et al.* The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* **39**, 27–38 (1984).
- Hobbs, H. H., Russell, D. W., Brown, M. S. & Goldstein, J. L. The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu. Rev. Genet.* **24**, 133–170 (1990).
- Juvonen, E., Ikkala, E., Fyhrquist, F. & Ruutu, T. Autosomal dominant erythrocytosis caused by increased sensitivity to erythropoietin. *Blood* **78**, 3066–3069 (1991).
- Doray, B. *et al.* Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease. *Hum. Mol. Genet.* **7**, 1449–1452 (1998).
- Glazier, A. M., Nadeau, J. H. & Aitman, T. J. Finding genes that underlie complex traits. *Science* **298**, 2345–2349 (2002).
- Kent, W. J. The BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
- Solovyyev, V. V. in *Bioinformatics From Genomes To Drugs. V.1. Basic Technologies* (ed. Lengauer, T.) 59–111 (Wiley-VCH, Weinheim, 2002).
- Birney, E. & Durbin, R. Using GeneWise in the *Drosophila* annotation experiment. *Genome Res.* **10**, 547–548 (2002).
- Lewis, S. E. *et al.* Apollo: a sequence annotation editor. *Genome Biol.* **3**, research0082.1–0082.14 (2002).
- Schwartz, S. *et al.* Human-mouse alignments with BLASTZ. *Genome Res.* **13**, 103–107 (2003).
- Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J. & Eichler, E. E. Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res.* **11**, 1005–1017 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We are grateful to our colleagues, S. Scherer, L. French and The Whitehead Institute/MIT Center for Genome Research for assistance with the screening of map gaps. We also acknowledge the HUGO Gene Nomenclature Committee for determining chromosome 19 gene symbols, H. Riethman for guidance regarding the telomeric regions of the chromosome, and U. Francke for helping to compile Supplementary Table S7 as well as for useful discussions regarding disease gene locations. This work was performed under the auspices of the US DOE's Office of Science, Biological and Environmental Research Program, by the University of California, Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, and Stanford University.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.G. (jane@shgc.stanford.edu) and E.M.R. (emrubin@lbl.gov). The entire sequence for the chromosome is deposited in GenBank under accession numbers NT_00255, NT_077812, NT_011295 and NT_011109.

Determination of electron orbital magnetic moments in carbon nanotubes

E. D. Minot*, Yuval Yaish*, Vera Sazonova & Paul L. McEuen

Laboratory of Atomic and Solid-State Physics, Cornell University, Ithaca, New York 14853, USA

* These authors contributed equally to this work

The remarkable transport properties of carbon nanotubes (CNTs) are determined by their unusual electronic structure¹. The electronic states of a carbon nanotube form one-dimensional electron and hole sub-bands, which, in general, are separated by an energy gap^{2,3}. States near the energy gap are predicted^{4,5} to have an orbital magnetic moment, μ_{orb} , that is much larger than the Bohr magneton (the magnetic moment of an electron due to its spin). This large moment is due to the motion of electrons around the circumference of the nanotube, and is thought to play a role in the magnetic susceptibility of CNTs⁶⁻⁹ and the magnetoresistance observed in large multiwalled CNTs¹⁰⁻¹². But the coupling between magnetic field and the electronic states of individual nanotubes remains to be quantified experimentally. Here we report electrical measurements of relatively small diameter (2–5 nm) individual CNTs in the presence of an axial magnetic field. We observe field-induced energy shifts of electronic states and the associated changes in sub-band structure, which enable us to confirm quantitatively the predicted values for μ_{orb} .

The electronic structure of a CNT is elegantly described by the quantization of wave states around a graphene cylinder¹. Graphene is a zero-bandgap semiconductor in which the valence and conduction states meet at two points in k -space, K_1 and K_2 (Fig. 1a). The dispersion around each of these points is a cone (Fig. 1b). When graphene is wrapped into a cylinder the electron wavenumber perpendicular to the CNT axis, $k_{\perp}$, is quantized, satisfying the boundary condition $\pi D k_{\perp} = 2\pi j$ where D is the CNT diameter and j is an integer. The resulting allowed values of k correspond to the horizontal lines in Fig. 1a that miss K_i by an amount $\Delta k_{\perp}$. The conic

sections of the dispersion cones by allowed k determine the CNT band structure near the Fermi level as shown in Fig. 1b. The upper and lower branches of the conic sections correspond to the conduction and valence states of the CNT. Both the K_1 and K_2 sub-bands have the same energy gap between conduction and valence states: $E_g^0 = \hbar v_F \Delta k_{\perp}$.

The size of $\Delta k_{\perp}$, and therefore E_g^0 , depends on the CNT chirality¹⁻³ and perturbations such as curvature¹³, axial strain^{14,15}, twist¹⁴ and inner–outer shell interactions¹⁶. From consideration of chirality alone, CNTs are classified as metallic ($\Delta k_{\perp} = 0$) or semiconducting ($\Delta k_{\perp} = 2/3D$)¹. Perturbations displace the dispersion cones¹³, modifying $\Delta k_{\perp}$ and resulting in an important class of small-bandgap ‘quasi-metallic’ CNTs¹⁷. We have used these small-bandgap CNTs in our measurements.

The electron states near the gap correspond to semiclassical electron orbits encircling the CNT. The perpendicular component of orbital velocity $v_{\perp} = (1/\hbar) dE/dk_{\perp}$ determines the clockwise (CW) or anticlockwise (ACW) sense of an orbit. For example, in Fig. 1b we see that $v_{\perp}$ is negative for the K_1 conduction states but is positive for K_1 valence states. By symmetry, each CW (ACW) orbit in the K_1 sub-band has an equal energy ACW (CW) partner in the K_2 sub-band. As a consequence, the two sub-bands are degenerate, but the CW/ACW sense of valence and conduction states is reversed.

From basic electromagnetic theory, an electron moving at velocity v around a loop of diameter D has an orbital magnetic moment of magnitude $\mu = Dev/4$. In a CNT, electron states at the bandgap edges, where $v_{\perp}$ is largest, have an orbital magnetic moment of magnitude $\mu_{\text{orb}} = Dev_F/4$ directed along the tube axis. A magnetic field parallel to the CNT axis, $B_{\parallel}$, is predicted to shift the energy of these states by:

$$\Delta E = -\mu_{\text{orb}} \cdot \mathbf{B} = \pm \frac{Dev_F B_{\parallel}}{4} \quad (1)$$

For CNTs with a finite energy gap at $B_{\parallel} = 0$, the energy gap of one sub-band becomes larger as $B_{\parallel}$ is increased, while the energy gap of the other sub-band becomes smaller (Fig. 1c).

Previous work on the magnetoresistance of individual multiwalled nanotubes¹⁰⁻¹² and the magnetic susceptibility of CNT mats^{6,7} has not confirmed the magnitude of μ_{orb} or the splitting of sub-band degeneracy. In the current work we use two different techniques to achieve this goal: (1) thermally activated transport through individual small-bandgap CNTs that are depleted of charge

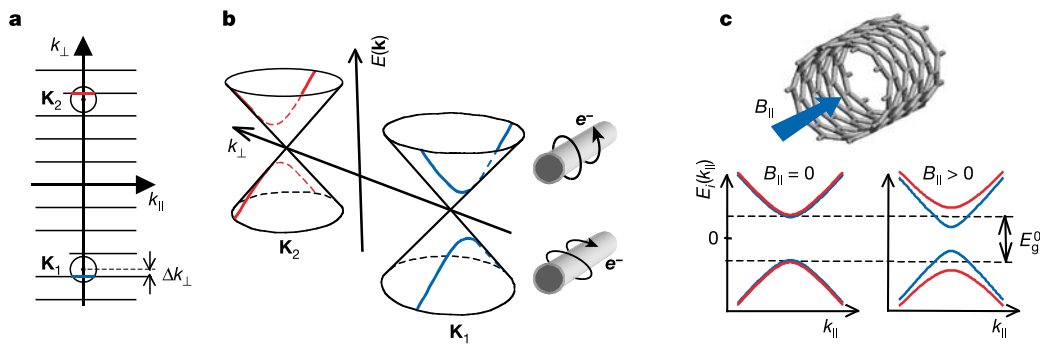


Figure 1 Nanotube states near the bandgap and orbital magnetic moments. **a**, The valence and conduction states of graphene meet at K_1 and K_2 . Horizontal lines show the quantized values of $k_{\perp}$ for the CNT structure in **c**. The misalignment between horizontal lines and the K -points is $\Delta k_{\perp}$. **b**, Graphene dispersion near the K -points is described by the cones $E_i(\mathbf{k}) = \pm \hbar v_F |\mathbf{k} - \mathbf{K}_i|$, with $v_F = 8 \times 10^5 \text{ m s}^{-1}$ (ref. 1). Lines of allowed k intersect the two cones (blue and red curves). The conduction states near K_1 (upper blue curve) have $dE/dk_{\perp} < 0$. Electrons in these states move around the CNT in an anticlockwise (ACW) fashion. The valence states near K_1 (lower

blue curve) have $dE/dk_{\perp} > 0$ and are associated with clockwise (CW) electron motion. ACW (CW) orbits correspond to positive (negative) magnetic moments along the CNT axis. The conic section near K_2 lies on the opposite face of an identical dispersion cone. Therefore, K_2 conduction (valence) states have CW (ACW) orbits. **c**, Top, perspective view of a CNT in the presence of a magnetic field $B_{\parallel}$. Below, the dispersion relations $E_1(k_{\parallel})$ and $E_2(k_{\parallel})$, shown in blue and red respectively. The sub-bands are degenerate at $B_{\parallel} = 0$. The magnetic field breaks this degeneracy.

carriers, (2) energy level spectroscopy near the bandgap edge of CNT quantum dots.

We have found that a suspended CNT device geometry (Fig. 2) is well suited for studying small changes in bandgap. Measurements of many such devices, using a gold-coated atomic force microscope (AFM) tip as a movable, local electrode^{18,19}, show that CNT segments contacting the oxide substrate are doped p-type, while suspended sections of the same tube are almost intrinsic. At small gate voltage V_g the suspended section is depleted of charge carriers. The oxide-bound sections, however, remain p-doped and act as electrodes to the suspended section. By studying the conductance of the suspended section at different temperatures and magnetic fields we can determine changes in $E_g^{K_i}$.

Figure 3a shows device conductance G versus V_g of two small-bandgap CNTs. Device 1 shows a sharp dip near $V_g = 0.4$ V, corresponding to depletion of carriers in the suspended segment. A second, broader dip occurs at $V_g \approx 2$ V as the oxide-bound segments become depleted. The inset shows the dip from the suspended section of device 2. In both cases, the addition of a magnetic field substantially increases the conductance at the bottom of the dip.

When the suspended CNT segment is depleted, conductance occurs via thermal activation of carriers across the energy gap. Conductance is smallest at $V_g = V^*$, immediately before the suspended segment becomes n-type (Fig. 2c). The minimum conductance due to thermal activation, $G_{\text{act}}(V^*)$, can be estimated by considering the Fermi-Dirac function at temperature T and the Landauer formalism for one-dimensional (1D) conduction channels^{15,20}

$$G_{\text{act}}(V^*, T) = \frac{2e^2}{h} \sum_{i=1,2} |t_i|^2 \frac{2}{\exp(E_g^{K_i}/k_B T) + 1} \quad (2)$$

where $|t_i|^2$ is the transmission probability for thermally activated carriers in the i th sub-band. The device conductance G is a combination of G_{act} in series with the conductance of the p-type sections of CNT and the conductance of the metal–CNT contacts, both of which are largely temperature independent.

We have measured G versus V_g for devices 1 and 2 at several temperatures. In Fig. 3b (open circles) we plot the change in resistance $\Delta R(T) = G(V^*, T)^{-1} - G(V_g \ll 0, T)^{-1}$ of device 2 at $B = 0$ T. From the slope and intercept of the fitting exponential, and assuming sub-band degeneracy ($E_g^{K_i} = E_g^0$), we find: $E_g^0 = 40$ meV and $|t_1|^2 + |t_2|^2 = 1.6$. Because $|t_1|^2 + |t_2|^2$ is close to 2, we conclude that transport is nearly ballistic and that both the K_1 and K_2 sub-bands make comparable contributions to the device conductance. We find similar values of E_g^0 in both devices (see Table 1) even though the CNT diameters are significantly different. This suggests that the bandgaps are not curvature related¹³. Further work is needed to identify the perturbations responsible for E_g^0 .

Magnetic fields dramatically reduce ΔR , as shown in Fig. 3c. The temperature dependence of ΔR at $B = 10$ T is also shown for device 2 (Fig. 3b, filled triangles). If we fit this high-field temperature data with the same method used for zero-field data, we find $E_g^0 = 22$ meV and $|t_1|^2 + |t_2|^2 = 0.8$. The bandgap of at least one sub-band is significantly lowered by the magnetic field and we argue below that

the apparent change in $|t_1|^2 + |t_2|^2$ is due to the increasing bandgap of the second sub-band.

The magnetic field dependence of ΔR can be quantitatively described by equal and opposite changes in $E_g^{K_1}$ and $E_g^{K_2}$ owing to the coupling of μ_{orb} with $B_{\parallel}$. We have accurately fitted our measurements of $\Delta R(B, T)$ using equation (2) and setting $E_g^{K_1} = E_g^0 - aB$ and $E_g^{K_2} = E_g^0 + aB$ (see the fitted curves in Fig. 3c). The only fitting parameter is a ; E_g^0 and $|t_i|^2$ are found from the temperature dependence of ΔR at $B = 0$ T and setting $|t_1|^2 = |t_2|^2$.

The fitting results for devices 1 and 2 are summarized in Table 1. In agreement with equation (1), the measured μ_{orb} scale with diameter and are an order of magnitude larger than previously measured spin magnetic moments in CNTs^{21,22}. Thermally activated transport (equation (2)), combined with the breaking of CW/ACW sub-band degeneracy, describes ΔR over a wide range of T and B . At $B = 10$ T device conductance is almost entirely due to carriers that are thermally activated across the smaller bandgap. Transport occurs in a single sub-band, explaining why $|t_1|^2 + |t_2|^2$ decreases by a factor of 2 when sub-band degeneracy is incorrectly assumed at high field. Our measurements confirm theoretical predictions^{4,5} for the sign and magnitude of orbital magnetic moments in CNTs and show that an applied magnetic field can split the degeneracy of the K_1 and K_2 sub-bands.

Orbital magnetic moments should also influence the energy level spectra of CNT quantum dots (CNTQDs) in applied magnetic fields. In our device geometry a CNTQD forms when $V_g > V^*$ and electrons are confined to conduction states of the suspended section by p–n tunnel barriers (Fig. 2c). Figure 4a shows the formation of a

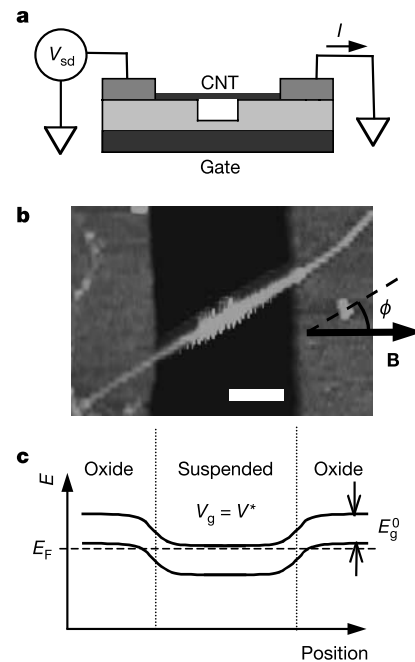


Figure 2 Device geometry and band bending. **a**, CNTs are grown on Si/SiO_x substrates by the chemical vapour deposition method²⁶. Electrodes (5 nm Cr, 50 nm Au) are patterned by photolithography²⁷. The central region of the CNT is suspended over a trench defined by electron-beam lithography and wet etching using 6:1 buffered HF. **b**, AFM image of the suspended section of CNT and nearby oxide-bound sections of device 1. Scale bar, 130 nm. The suspended section appears fuzzy because it is displaced by the AFM tip during imaging. From the image we find CNT diameter $D = 2.6$ nm, suspended length $L = 500$ nm, and determine the misalignment angle ϕ between applied magnetic field and the CNT axis. **c**, Band bending in the suspended CNT segment and neighbouring oxide-bound segments when $V_g = V^*$. The number of thermally activated carriers is minimized and there is no n-type region to facilitate tunnelling processes. The oxide-bound sections remain p-type at small V_g .

Table 1 Summary of thermal activation results

	D (nm)	E_g^0 (meV)	ϕ (°)	a (meV T ⁻¹)	μ_{orb} (meV T ⁻¹)	
					Experiment	Theory
Device 1	2.6 ± 0.3	36 ± 3	30 ± 3	1.3 ± 0.1	0.7 ± 0.1	0.5 ± 0.1
			60 ± 3	0.7 ± 0.1	0.7 ± 0.1	0.5 ± 0.1
Device 2	5.0 ± 0.3	40 ± 3	45 ± 3	2.1 ± 0.2	1.5 ± 0.2	1.0 ± 0.2

ϕ is the misalignment angle between CNT axis and the magnetic field direction. The experimental value of μ_{orb} is given by $a/2\cos\phi$. There is uncertainty in theoretical values of μ_{orb} owing to uncertainty in V_F and D .

CNTQD in device 1 at $V_g > V^*$, $T = 1.5$ K. There is a large region of zero conductance as the Fermi level passes through the energy gap of the suspended section. At higher V_g the Coulomb diamonds labelled 1, 2, 3 and 4 correspond to charge states of one, two, three and four electrons in the conduction band of the suspended segment.

In the Coulomb blockade model of quantum dots²³, the width of the N th diamond is proportional to a fixed electrostatic charging energy plus the energy difference between the quantum levels occupied by N th and $(N + 1)$ th electrons. The energies of the quantum levels in our CNTQD can be estimated by considering electrons confined to a 1D potential well of length L . The confinement results in quantized $k_{||}$ values which, combined with the dispersion relations $E_i(k_{||})$, determine the energy levels of the dot. Near the bandgap edge $E_i(k_{||})$ are parabolic, therefore, the energy levels of the first few conduction states should be:

$$\varepsilon(n, i, B_{||}) = \frac{E_g^0}{2} + \frac{\hbar^2 \pi^2}{2m_i^* L^2} n^2 \pm \mu_{orb} B_{||} \quad (3)$$

where the quantum number n is a positive integer, the effective mass $m_i^* = E_{g,i}^0 / 2v_F^2$, and $+$ applies to CW orbitals while $-$ applies to ACW orbitals. The first few level crossings predicted by equation (3) are shown in Fig. 4b.

Figure 4c shows low-bias $G-V_g$ plots of a Coulomb peak from device 1 as B is increased. The peak corresponds to the second electron added to the dot (the intersection of Coulomb diamonds 1 and 2). The peak shifts ~ 1.2 mV T^{-1} and doubles in conductance as B reaches 3.6 T. Figure 4d shows the first eight Coulomb peaks. The positions of the peaks generally move between 1.2 and 1.6 mV T^{-1} . The fifth and subsequent peaks show clear changes between positive and negative slopes. Peaks appear to be paired, each pair having a different zigzag pattern.

The main features of Fig. 4d are described by the CNTQD model. Peaks with $d\varepsilon/dB > 0$ correspond to tunnelling into a CW orbital, while peaks with $d\varepsilon/dB < 0$ correspond to tunnelling into a ACW orbital. The measured value of $\mu_{orb} = |d\varepsilon/dB| = 0.7 \pm$

0.1 meV T^{-1} is inferred as described in the legend of Fig. 4, and agrees with the values in Table 1 for device 1. Furthermore, the striking difference between the first four peaks and later peaks is in qualitative agreement with the modelled spectrum (Fig. 4b). The first pair of peaks (spin up and spin down, $n = 1$, ACW orbital) are not expected to undergo level crossings. The second pair (peaks 3 and 4 in Fig. 4d) may undergo a level crossing at low field, but the resolution of our data is limited by thermal broadening; levels separated by less than $4k_B T \approx 0.5$ meV merge together. The third and fourth pairs clearly show the changes in slope that are expected when level crossings occur. We conclude that there are quantum levels near the bandgap edges with both positive and negative orbital magnetic moments whose magnitudes are consistent with theoretical predictions^{4,5}. The Coulomb blockade model does not describe all the features in Fig. 4d. The detailed structure of this CNTQD system may depend on effects such as exchange coupling^{21,24}, and will be the subject of future work.

Our measured values of μ_{orb} are 10–20 times larger than the Bohr

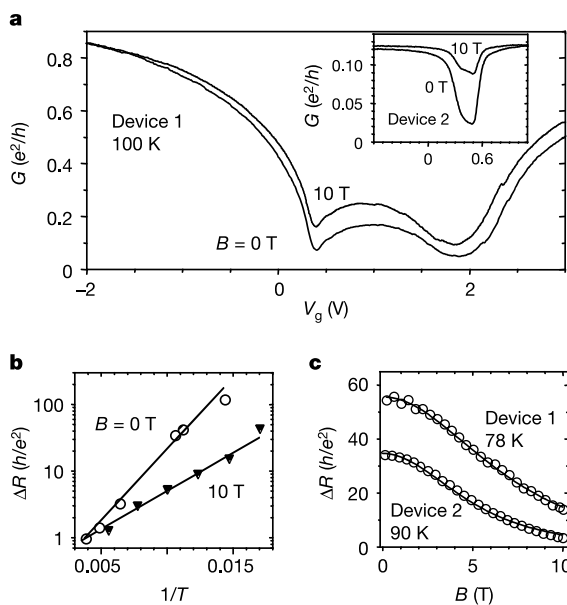


Figure 3 Effect of magnetic field on device resistance. **a**, $I-V_g$ curves for devices 1 and 2 at $T = 100$ K. Curves taken at $B = 0$ T have lower conductance than curves taken at $B = 10$ T. **b**, ΔR as a function of $1/T$ for device 2. The data shown are for $B = 0$ T (larger ΔR) and $B = 10$ T (smaller ΔR). **c**, ΔR as a function of B for device 1 at $T = 78$ K (upper curve) and device 2 at $T = 90$ K (lower curve).

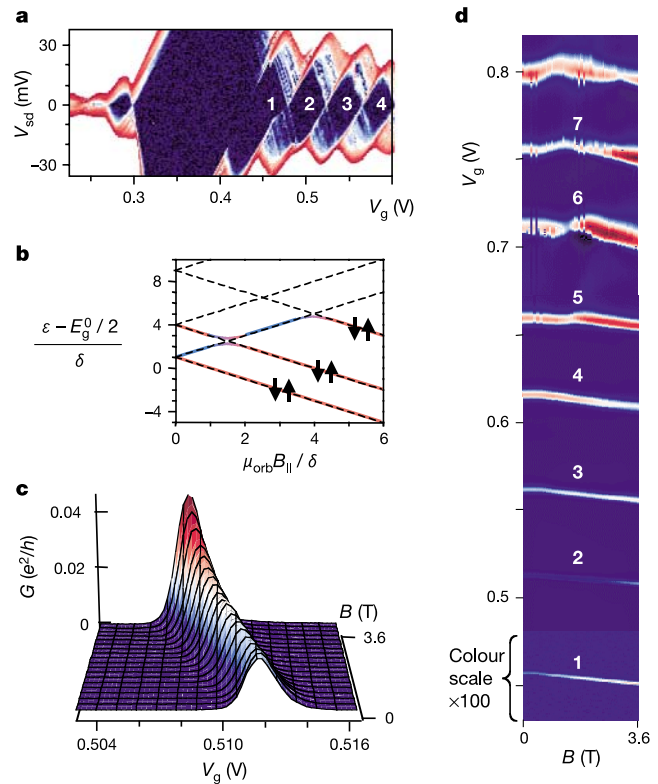


Figure 4 Energy levels of a nanotube quantum dot. **a**, Differential conductance dI/dV_{sd} as a function of source-drain voltage, V_{sd} , and V_g . Data are from device 1 at $T = 1.5$ K. Dark blue represents $dI/dV_{sd} = 0$, dark red represents $dI/dV_{sd} = 0.2 e^2/h$. In the white regions (top and bottom of the plot), current levels exceeded the measurement range. The first four Coulomb diamonds, corresponding to discrete charge states, are labelled 1–4. The gate coupling α is twice the ratio of Coulomb diamond width to Coulomb diamond height²³. For this device $\alpha \approx 2.2$. **b**, Modelled energies of quantum levels from equation (3), approximating m_i^* as constant. The energy scale $\delta = \hbar^2 \pi^2 / 2m_i^* L^2$. For device 1 we have $\delta \approx 0.25$ meV. Coloured lines represent expected zigzags in the first six Coulomb peaks, with red and blue representing respectively ACW and CW states. Arrows indicate spin degeneracy for each state. **c**, Conductance I/V_{sd} as a function of magnetic field B for the second Coulomb peak of device 1, $\phi = 30^\circ$, $V_{sd} = 0.5$ mV. Shifts in peak position $V_g^0(n, i)$ are related to energy shifts of quantum levels by: $dV_g^0(n, i)/dB = \alpha d\varepsilon(n, i)/dB$. **d**, Low-bias conductance I/V_{sd} as a function of V_g and B showing first eight Coulomb peaks of device 1, $\phi = 30^\circ$. Dark blue represents $I/V_{sd} = 0$; dark red represents $I/V_{sd} = 0.35 e^2/h$. The colour scale for peak 1 is magnified by 100 times.

magneton and the spin magnetic moment in CNTs^{21,22}. The reason is the large size of electron orbits encircling the CNT compared to the radii of atomic orbitals. These large magnetic moments give researchers a powerful new tool to control the energy structure of CNTs. For example, the tunnel transparency of p–n barriers can be tuned by using a magnetic field to modify the bandgap. This effect is seen in Fig. 4c: the conductance of the Coulomb peak increases as the tunnel barriers become more transparent. This will be useful, for example, to study Kondo physics in CNTQDs^{24,25} at different tunnelling strengths. Researchers can also tune the energy levels of electrons in the 1D box formed by a CNT. By applying large magnetic fields it is possible to investigate the properties of a CNT in which only one sub-band is occupied. Conversely, by matching the energies of different sub-band states, the interactions between states arising from CW and ACW orbits can be explored. □

Received 18 December 2003; accepted 18 February 2004; doi:10.1038/nature02425.

- Sohn, L. L., Kouwenhoven, L. P. & Schon, G. (eds) *Carbon Nanotubes* (Springer, New York, 2001).
- Wildoer, J. W. G., Venema, L. C., Rinzler, A. G., Smalley, R. E. & Dekker, C. Electronic structure of atomically resolved carbon nanotubes. *Nature* **391**, 59–62 (1998).
- Odom, T. W., Huang, J. L., Kim, P. & Lieber, C. M. Atomic structure and electronic properties of single-walled carbon nanotubes. *Nature* **391**, 62–64 (1998).
- Ajiki, H. & Ando, T. Electronic states of carbon nanotubes. *J. Phys. Soc. Jpn* **62**, 1255–1266 (1993).
- Lu, J. P. Novel magnetic-properties of carbon nanotubes. *Phys. Rev. Lett.* **74**, 1123–1126 (1995).
- Ramirez, A. P. et al. Magnetic-susceptibility of molecular carbon—nanotubes and fullerite. *Science* **265**, 84–86 (1994).
- Wang, X. K., Chang, R. P. H., Patashinski, A. & Ketterson, J. B. Magnetic-susceptibility of buckytubes. *J. Mater. Res.* **9**, 1578–1582 (1994).
- Chauvet, O. et al. Magnetic anisotropies of aligned carbon nanotubes. *Phys. Rev. B* **52**, R6963–R6966 (1995).
- Walters, D. A. et al. In-plane-aligned membranes of carbon nanotubes. *Chem. Phys. Lett.* **338**, 14–20 (2001).
- Fujiwara, A., Tomiyama, K., Suematsu, H., Yumura, M. & Uchida, K. Quantum interference of electrons in multiwall carbon nanotubes. *Phys. Rev. B* **60**, 13492–13496 (1999).
- Bachtold, A. et al. Aharonov-Bohm oscillations in carbon nanotubes. *Nature* **397**, 673–675 (1999).
- Lee, J. O. et al. Observation of magnetic-field-modulated energy gap in carbon nanotubes. *Solid State Commun.* **115**, 467–471 (2000).
- Kane, C. L. & Mele, E. J. Size, shape, and low energy electronic structure of carbon nanotubes. *Phys. Rev. Lett.* **78**, 1932–1935 (1997).
- Yang, L. & Han, J. Electronic structure of deformed carbon nanotubes. *Phys. Rev. Lett.* **85**, 154–157 (2000).
- Minot, E. D. et al. Tuning carbon nanotube band gaps with strain. *Phys. Rev. Lett.* **90**, 156401 (2003).
- Kwon, Y. K. & Tomanek, D. Electronic and structural properties of multiwall carbon nanotubes. *Phys. Rev. B* **58**, R16001–R16004 (1998).
- Zhou, C. W., Kong, J. & Dai, H. J. Intrinsic electrical properties of individual single-walled carbon nanotubes with small band gaps. *Phys. Rev. Lett.* **84**, 5604–5607 (2000).
- de Pablo, P. J. et al. Nonlinear resistance versus length in single-walled carbon nanotubes. *Phys. Rev. Lett.* **88**, 036804 (2002).
- Yaish, Y. et al. Electrical nanoprobe of semiconducting carbon nanotubes using an atomic force microscope. *Phys. Rev. Lett.* **92**, 046401 (2004).
- Maiti, A., Svizhenko, A. & Anantram, M. P. Electronic transport through carbon nanotubes: Effects of structural deformation and tube chirality. *Phys. Rev. Lett.* **88**, 126805 (2002).
- Tans, S. J., Devoret, M. H., Groeneveld, R. J. A. & Dekker, C. Electron-electron correlations in carbon nanotubes. *Nature* **394**, 761–764 (1998).
- Cobden, D. H., Bockrath, M., McEuen, P. L., Rinzler, A. G. & Smalley, R. E. Spin splitting and even-odd effects in carbon nanotubes. *Phys. Rev. Lett.* **81**, 681–684 (1998).
- Sohn, L. L., Kouwenhoven, L. P. & Schon, G. (eds) *Mesoscopic Electron Transport* (Kluwer, Dordrecht, 1997).
- Liang, W. J., Bockrath, M. & Park, H. Shell filling and exchange coupling in metallic single-walled carbon nanotubes. *Phys. Rev. Lett.* **88**, 126801 (2002).
- Nygard, J., Cobden, D. H. & Lindelof, P. E. Kondo physics in carbon nanotubes. *Nature* **408**, 342–346 (2000).
- Kong, J., Soh, H. T., Cassell, A. M., Quate, C. F. & Dai, H. J. Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers. *Nature* **395**, 878–881 (1998).
- Rosenblatt, S. et al. High performance electrolyte gated carbon nanotube transistors. *Nano Lett.* **2**, 869–872 (2002).

Acknowledgements We thank H. Ustunel, T. Arias and H. Dai for discussions. This work was supported by the NSF through the Cornell Center for Materials Research and the Center for Nanoscale Systems, and by the MARCO Focused Research Center on Materials, Structures, and Devices. Sample fabrication was performed at the Cornell node of the National Nanofabrication Users Network, funded by NSF. One of us (E.D.M.) acknowledges support by an NSF graduate fellowship.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.L.M. (mceuen@ccmr.cornell.edu).

Current-induced domain-wall switching in a ferromagnetic semiconductor structure

M. Yamanouchi¹, D. Chiba¹, F. Matsukura^{1,2} & H. Ohno^{1,2}

¹Laboratory for Nanoelectronics and Spintronics, Research Institute of Electrical Communication, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

²ERATO Semiconductor Spintronics Project, Japan Science and Technology Agency, Japan

Magnetic information storage relies on external magnetic fields to encode logical bits through magnetization reversal. But because the magnetic fields needed to operate ultradense storage devices are too high to generate, magnetization reversal by electrical currents is attracting much interest as a promising alternative encoding method. Indeed, spin-polarized currents can reverse the magnetization direction of nanometre-sized metallic structures through torque^{1–4}; however, the high current densities of 10^7 – 10^8 A cm^{–2} that are at present required exceed the threshold values tolerated by the metal interconnects of integrated circuits^{5,6}. Encoding magnetic information in metallic systems has also been achieved by manipulating the domain walls at the boundary between regions with different magnetization directions^{7–13}, but the approach again requires high current densities of about 10^7 A cm^{–2}. Here we demonstrate that, in a ferromagnetic semiconductor structure, magnetization reversal through domain-wall switching can be induced in the absence of a magnetic field using current pulses with densities below 10^5 A cm^{–2}. The slow switching speed and low ferromagnetic transition temperature of our current system are impractical. But provided these problems can be addressed, magnetic reversal through electric pulses with reduced current densities could provide a route to magnetic information storage applications.

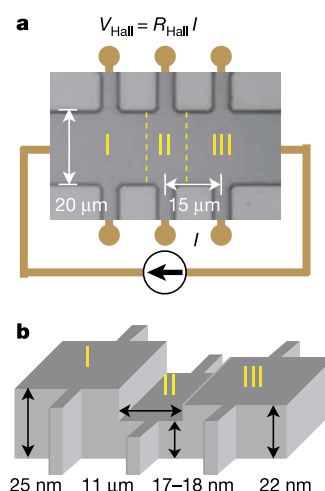


Figure 1 A micrograph and a schematic drawing of the device. **a**, A 20-μm-wide channel with three pairs of Hall probes separated by 15 μm was defined by photolithography and wet etching. **b**, To reduce the coercive force of the two regions, 7–8 nm and 3 nm of the surface layers of regions II and III, respectively, were removed. A domain wall was prepared at the boundary of regions I and II, and its position after application of a current pulse was monitored by $R_{\text{Hall}} = V_{\text{Hall}}/I$, using a small probe current I . Hall voltage V_{Hall} is measured at the Hall probes. Magneto-optical Kerr microscopy was also used to image the domain structure.

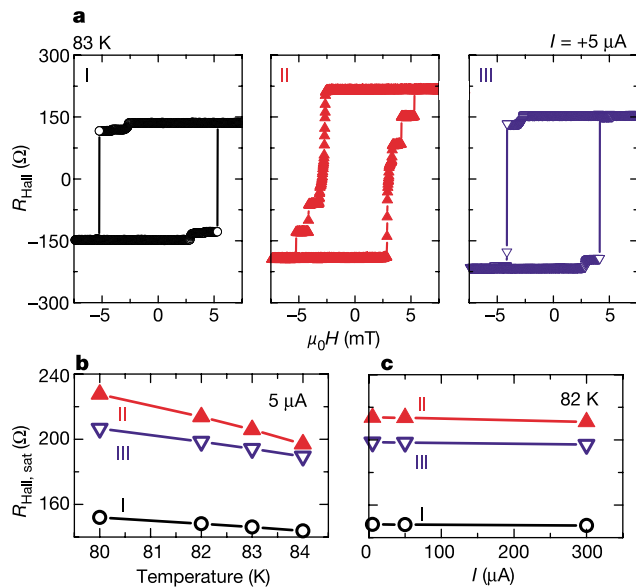


Figure 2 The hysteresis loops of regions I, II and III, of sample A measured by R_{Hall} at 83 K, and the temperature and the current dependence of the averaged saturated Hall resistance, $R_{\text{Hall,sat}}$. **a**, The coercive force $\mu_0 H_c$, the applied magnetic field at $R_{\text{Hall}} = 0$, shows that $H_c(\text{I}) > H_c(\text{III}) > H_c(\text{II})$, as designed. No dependence of hysteresis on current direction was observed at $I = 5 \mu\text{A}$. The difference in $R_{\text{Hall,sat}}$ reflects the difference in the anomalous Hall coefficient due to resistance difference. Small steps in the hysteresis are caused by the magnetization rotation of adjacent region(s). **b**, Temperature dependence of $R_{\text{Hall,sat}}$ measured using 5- μA , 1.5-s current pulses; open circles indicate $R_{\text{Hall,sat}}$ for region I, closed triangles for II, and open downward triangles for III. **c**, Current dependence of $R_{\text{Hall,sat}}$ measured using 1.5-s current pulses at 82 K. Symbols are the same as those in **b**. Comparison of **b** and **c** shows that the temperature increase is less than 0.4 K with 300- μA current pulses.

Figure 1 shows the device structure used for the current-induced domain-wall switching. The semiconductor used is an alloy of Mn and GaAs that exhibits carrier-induced ferromagnetism^{14–16}: (Ga,Mn)As. The starting single-crystal multilayer was grown by molecular beam epitaxy¹⁴ with the layer structure from the surface side; 25 nm $(\text{Ga}_{1-x}\text{Mn}_x)\text{As}/500 \text{ nm } (\text{In}_y\text{Ga}_{1-y})\text{As}/100 \text{ nm GaAs}$ on (001) semi-insulating GaAs substrate. Two samples, A (for which $x = 0.050$, $y = 0.15$) and B (for which $x = 0.038$, $y = 0.23$), are discussed here. The ferromagnetic transition temperature T_c of the (Ga,Mn)As layer is 90 K for sample A and 110 K for sample B. The (In,Ga)As buffer layer introduces tensile strain in the (Ga,Mn)As layer so that the magnetic easy axis lies along the growth direction¹⁶. This allows the measurement of hysteresis loops and domain-wall switching through the Hall resistance R_{Hall} , which is dominated by the plane component of magnetization M of the ferromagnetic semiconductor layer.

A 20- μm -wide channel along the $\langle 110 \rangle$ direction was defined using conventional photolithography and wet etching (Fig. 1a). The channel consists of three regions—I, II and III—having different (Ga,Mn)As thicknesses: the surfaces of regions II and III were removed for 7–8 nm and 3 nm, respectively, with the length of region II being 11 μm (Fig. 1b). This structure makes it possible to prepare a domain wall at the boundary of I and II, because the removal of the surface layer results in a change of the coercive force H_c , at which magnetization reversal takes place, through a change of T_c (ref. 17). It also confines the domain wall inside region II, because the step at each boundary acts as a barrier for the domain-wall motion¹⁸. A pair of Hall probes was attached to each of the three regions to monitor M of the respective portion of the device. A

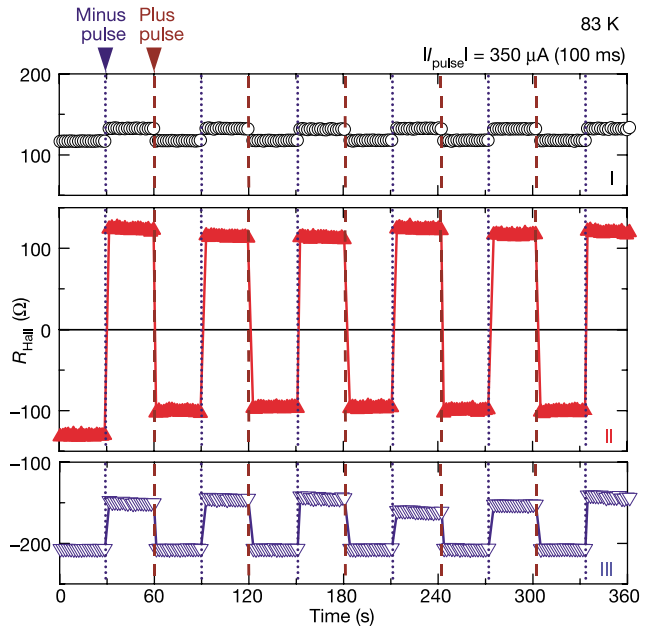


Figure 3 The effect of successive alternating negative and positive current pulses on R_{Hall} for regions I, II and III of sample A. The amplitude of the pulse is 350 μA and its width 100 ms. The dotted vertical lines indicate time t at which a negative current pulse is applied, whereas the dashed vertical lines indicate time t at which a positive pulse is applied. The current direction from regions I to III is defined to be positive. No magnetic field is applied. At $t = 0$, the magnetization M of region I is antiparallel to M of regions II and III because the sign of R_{Hall} in the respective regions is opposite: that is, a domain wall is at the boundary of regions I and II (see text for the preparation of the initial state). When a negative current pulse is applied ($t = 30 \text{ s}$), R_{Hall} of region II switches its sign, indicating that the M direction of region II is reversed by the current pulse. The M direction of region II can be switched back, as shown by the sign change in R_{Hall} of region II by a positive current pulse ($t = 60 \text{ s}$). This current switching of the domain wall—that is, the M reversal of region II—can be repeated, as shown by the successive sign changes. The reproducibility of R_{Hall} values of each region is a measure of that of the position of the domain wall.

magneto-optical polar-Kerr-effect (MOKE) microscope was used to image the domain configurations.

The Hall resistances R_{Hall} of sample A as a function of applied magnetic field $\mu_0 H$ (where μ_0 is the permeability of vacuum) shown in Fig. 2a reveal the magnetization hysteresis loops of regions I, II and III at 83 K. As we designed them to be, the order of the coercive forces is $H_c(\text{I}) > H_c(\text{III}) > H_c(\text{II})$. The saturated Hall resistance $R_{\text{Hall,sat}}$ reflects the resistance difference, because the anomalous Hall coefficient is a function of the channel resistance, and the small stepped structures in the hysteresis loops are caused by the magnetization rotation of the adjacent region(s). Because virtually the same hysteresis loop is obtained, regardless of the direction of the current at 5 μA , this current level was used to measure the effect of higher current levels on the domain-wall position. We also examined the effect of increasing current level on the sample temperature. Figure 2b shows the temperature dependence of $R_{\text{Hall,sat}}$ (see also Fig. 2a) and Fig. 2c shows $R_{\text{Hall,sat}}$ versus current, both measured using a pulse width of 1.5 s. A comparison of Fig. 2b and c reveals that the temperature increase is less than 0.4 K under application of 300- μA current pulses.

For the switching experiment, we first prepared an antiparallel magnetization configuration between regions I and II: we saturated the magnetization M of all three regions at 83 K by using a sufficiently high external positive field $\mu_0 H$, then reduced $\mu_0 H$ to -4.9 mT (which reverses the M of regions II and III but not of I) and

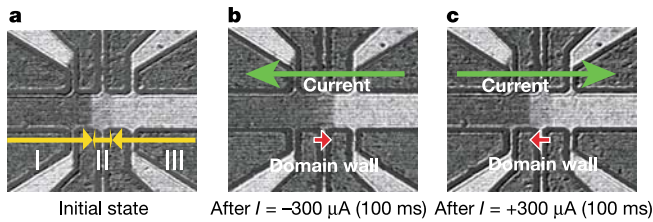


Figure 4 MOKE images of sample A using 546-nm light at ~ 80 K. Black and white regions in the channel correspond to positive and negative values of M , respectively. **a**, The MOKE image of the initial state, where the domain wall is at the left edge of region II. Regions I, II and III are indicated by arrows in the image. **b**, The MOKE image after application of a current pulse $I = -300 \mu\text{A}$ (100 ms), showing that the domain wall is now at the right edge of region II. **c**, A positive current pulse of $I = +300 \mu\text{A}$ (100 ms) switches the domain wall back to its original position.

finally removed the field so that $\mu_0 H = 0$. This initial condition ensures that the domain wall is at the boundary of regions I and II. R_{Hall} values at time $t = 0$ in Fig. 3, measured under the initial condition with a $5\text{-}\mu\text{A}$ current, show that the M values of regions I and II + III are indeed antiparallel, because the R_{Hall} in region I, $R_{\text{Hall}}(\text{I}) > 0$, whereas $R_{\text{Hall}}(\text{II})$ and $R_{\text{Hall}}(\text{III}) < 0$. At $t = 30$ s, a current pulse of $-350 \mu\text{A}$ with width 100 ms is applied: positive current is defined as a current flow from regions I to III. The R_{Hall} traces after the current pulse indicate that the domain wall was moved to the other end of region II by the pulse, resulting in $R_{\text{Hall}}(\text{I})$ and $R_{\text{Hall}}(\text{II}) > 0$, and $R_{\text{Hall}}(\text{III}) < 0$. The domain wall can be switched back to the other boundary by a positive current pulse of $350 \mu\text{A}$ for 100 ms, as is shown in the traces after the pulse at $t = 60$ s. This current-induced domain-wall switching in the absence of external magnetic field can be repeated, as was unambiguously detected by the successive sign change of R_{Hall} in region II shown in Fig. 3. The good reproducibility of R_{Hall} values is a measure of the reproducibility of the position of the domain wall after the application of each current pulse.

The current-induced domain-wall switching in sample A can be visualized using MOKE spectroscopy, as shown in Fig. 4. Figure 4a is an MOKE image of the initial domain configuration at about 80 K under a field of $\mu_0 H = 0$ prepared following the procedure described above. The bright region of the channel shows that M is pointing into the plane, whereas in the dark region M is pointing out of the plane. The centre part of the channel with reduced brightness indicates region II, where the thickness difference and associated change in the magnetic properties are reflected in the MOKE contrast. After application of a $-300\text{-}\mu\text{A}$, 100-ms current pulse, the MOKE image changes to the one shown in Fig. 4b. The domain wall is now at the right end of region II, showing that the current-induced domain-wall motion is in accordance with the Hall measurements. The domain wall can be switched back to its original position with a subsequent $+300\text{-}\mu\text{A}$, 100-ms current pulse (Fig. 4c). The nanometre scale steps at the two ends of region II act as a domain-wall stop, providing a way to control the domain-wall position precisely.

We note that the current density of the pulses, $8 \times 10^4 \text{ A cm}^{-2}$, is two to three orders of magnitude lower than that required in the metallic systems. Current density and pulse-width dependence experiments done on sample B (not shown) at temperature $T = 86\text{--}94$ K revealed that there exists a threshold current density (at around 10^4 A cm^{-2}) below which the domain wall does not appear to move. The speed of the wall defined by the channel length ($11 \mu\text{m}$), divided by the pulse width, reaches 0.2 ms^{-1} at $2 \times 10^5 \text{ A cm}^{-2}$.

We now compare the present results with the mechanisms proposed for the current-induced domain-wall motion in metallic systems. The influence of Joule heating is not important here,

because the sample temperature is virtually unchanged, as shown in Fig. 2c, and, more importantly, the effect depends on the direction of the current pulse. The magnetic field generated by the current pulse is of the order of 0.01 mT , too small compared to the $\mu_0 H_c$ of region II, and cannot be responsible for the observed effect. Domain-wall distortion by the leading and trailing edges of current pulses¹⁹ can also be excluded, because here the wall motion depends on the pulse width and the domain wall continuously moves at low d.c. current (not shown). The inhomogeneous current distribution in the vicinity of the domain wall due to the absence of Hall voltage²⁰ leads to a wall motion that is opposite to the present results because the Hall coefficient of (Ga,Mn)As is positive.

Although it is rather early to draw conclusions, the spin angular momentum transfer through p - d exchange interaction between localized Mn spin and itinerant hole^{21,22} appears to be a possible candidate for the mechanism behind the observed results. We point out that the sign of p - d exchange in (Ga,Mn)As is known to be negative¹⁶, which, according to this mechanism, should result in a domain-wall motion that is opposite to the current direction—consistent with our observations. If the angular momentum transfer process alone is responsible for the wall motion, the efficiency is as high as 8%, comparable to metallic systems¹³: the efficiency is the ratio of the change in the magnetic moment of the Mn spin system to that of the total number of holes (carriers) traversed during the pulse (where the parameters used to calculate the ratio are taken from the mean-field theory¹⁶), and suggests that a significant part of the carrier angular momentum is transferred, in spite of the presence of a strong spin-orbit interaction that mixes the two spin directions in p -type III-V ferromagnetic semiconductors.

The reduction of current density by two to three orders of magnitude for the magnetization reversal in a lithographically defined area without resorting to external magnetic fields is attractive for future applications. Because they can be seamlessly integrated into semiconductor structures, we may also expect ferromagnetic semiconductors to be useful in developing spintronic devices that combine the present switching with other effects such as the electric-field control of T_c and H_c (refs 23, 24) and the giant planar Hall effect²⁵, and device structures such as spin light-emitting diodes²⁶ and spin transistors²⁷. The recent reports on high- T_c ferromagnetic semiconductors are encouraging signs towards room-temperature observation of the present effect; a room-temperature T_c has been reported²⁸ and much higher T_c values have been claimed²⁹. □

Received 7 September 2003; accepted 27 February 2004; doi:10.1038/nature02441.

- Slonczewski, J. Current-driven excitation of magnetic multilayer. *J. Magn. Magn. Mater.* **159**, L1–L7 (1996).
- Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54**, 9353–9358 (1996).
- Tsoi, M. *et al.* Excitation of a magnetic multilayer by an electric current. *Phys. Rev. Lett.* **80**, 4281–4284 (1998) (erratum (81), 493 (1998)).
- Albert, F. J., Katine, J. A., Buhrman, R. A. & Ralph, D. C. Spin-polarized current switching of a Co thin film nanomagnet. *Appl. Phys. Lett.* **77**, 3809–3811 (2000).
- Sze, S. M. (ed.) *VLSI Technology* 412 (McGraw-Hill, New York, 1988).
- Michael, N. L., Kim, C. U., Gillespie, P. & Augur, R. in *Proc. Advanced Metallization Conf.*, 29–30 Oct. 2002, Tokyo (ed. Zaima, S.) 42–43 (Office Sofiel, Tokyo, 2002).
- Berger, L. Exchange interaction between ferromagnetic domain wall and electric current in very thin metallic films. *J. Appl. Phys.* **55**, 1954–1956 (1984).
- Freitas, P. P. & Berger, L. Observation of s-d exchange force between domain wall and electric current in very thin permalloy films. *J. Appl. Phys.* **57**, 1266–1269 (1985).
- Gan, L., Chung, S. H., Aschenbach, K. H., Dreyer, M. & Gomez, R. D. Pulsed-current-induced domain wall propagation in permalloy patterns observed using magnetic force microscope. *IEEE Trans. Mag.* **36**, 3047–3049 (2000).
- Grollier, J. *et al.* Switching a spin valve back and forth by current-induced domain wall motion. *Appl. Phys. Lett.* **83**, 509–511 (2003).
- Tsoi, M., Fontana, R. E. & Parkin, S. S. P. Magnetic domain wall motion triggered by an electric current. *Appl. Phys. Lett.* **83**, 2617–2619 (2003).
- Vernier, N., Allwood, D. A., Atkinson, D., Cooke, M. D. & Cowburn, R. P. Domain wall propagation in magnetic nanowires by spin polarized current injection. Preprint at (<http://lanl.arXiv.org/abs/cond-mat/0304549>) (2003).
- Yamaguchi, A. *et al.* Real-space observation of current-driven domain wall motion in submicron magnetic wires. *Phys. Rev. Lett.* **92**, 077205 (2004).
- Ohno, H. Making nonmagnetic semiconductors ferromagnetic. *Science* **281**, 951–956 (1998).

15. Dietl, T., Ohno, H., Matsukura, F., Cibert, J. & Ferrand, D. Zener model description of ferromagnetism in zinc-blende magnetic semiconductors. *Science* **287**, 1019–1022 (2000).
16. Dietl, T., Ohno, H. & Matsukura, F. Hole-mediated ferromagnetism in tetrahedrally coordinated semiconductors. *Phys. Rev. B* **63**, 195205 (2001).
17. Chiba, D., Takamura, K., Matsukura, F. & Ohno, H. Effect of low-temperature annealing on (Ga,Mn)As trilayer structures. *Appl. Phys. Lett.* **82**, 3020–3022 (2003).
18. Labaye, Y., Berger, L. & Coey, J. M. D. Domain walls in ferromagnetic nanoconstriction. *J. Appl. Phys.* **91**, 5341–5346 (2002).
19. Salhi, E. & Berger, L. Current-induced displacements of Bloch walls in Ni-Fe films of thickness 120–740 nm. *J. Appl. Phys.* **76**, 4787–4792 (1994).
20. Berger, L. Prediction of a domain-drag effect in uniaxial, non-compensated, ferromagnetic metals. *J. Phys. Chem. Solids* **35**, 947–956 (1974).
21. Berger, L. Motion of magnetic domain wall traversed by fast-rising current pulses. *J. Appl. Phys.* **71**, 2721–2726 (1992).
22. Tataru, G. & Kohno, H. Theory of current-driven domain wall motion: spin transfer versus momentum transfer. *Phys. Rev. Lett.* **92**, 086601 (2004).
23. Ohno, H. *et al.* Electric-field control of ferromagnetism. *Nature* **408**, 944–946 (2000).
24. Chiba, D., Yamanouchi, M., Matsukura, F. & Ohno, H. Electrical manipulation of magnetization reversal in a ferromagnetic semiconductor. *Science* **301**, 943–945 (2003).
25. Tang, H. X., Kawakami, R. K., Awschalom, D. D. & Roukes, M. L. Giant planar Hall effect in epitaxial (Ga,Mn)As devices. *Phys. Rev. Lett.* **90**, 107201 (2003).
26. Ohno, Y. *et al.* Electrical spin injection in a ferromagnetic semiconductor heterostructure. *Nature* **402**, 790–792 (1999).
27. Flatté, M. E., Yu, Z. G., Johnston-Halperin, E. & Awschalom, D. D. Theory of semiconductor magnetic bipolar transistors. *Appl. Phys. Lett.* **82**, 4740–4742 (2003).
28. Saito, H., Zayets, V., Yamagata, S. & Ando, K. Room-temperature ferromagnetism in a II–VI diluted magnetic semiconductor $\text{Zn}_{1-x}\text{Cr}_x\text{Te}$. *Phys. Rev. Lett.* **90**, 207202 (2003).
29. Pearton, S. J. *et al.* Wide band gap ferromagnetic semiconductors and oxides. *J. Appl. Phys.* **93**, 1–13 (2003).

Acknowledgements We thank T. Dietl, Y. Ohno, K. Ohtani and H. Kurino for discussions and K. Akahane, S. Meguro, S. Ito, K. Ikuno and Y. Sato for their technical assistance. This work was supported in part by the IT programme of ‘RR2002’ from MEXT, grant-in-aids from MEXT/SPS, research fellowships from JSPS, and the ‘21st Century COE’ programme at Tohoku University.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.O. (ohno@riec.tohoku.ac.jp).

Superconductivity in diamond

E. A. Ekimov¹, V. A. Sidorov¹, E. D. Bauer², N. N. Mel'nik³, N. J. Curro², J. D. Thompson² & S. M. Stishov¹

¹Vereshchagin Institute for High Pressure Physics, Russian Academy of Sciences, 142190 Troitsk, Moscow region, Russia

²Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

³Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

Diamond is an electrical insulator well known for its exceptional hardness. It also conducts heat even more effectively than copper, and can withstand very high electric fields¹. With these physical properties, diamond is attractive for electronic applications², particularly when charge carriers are introduced (by chemical doping) into the system. Boron has one less electron than carbon and, because of its small atomic radius, boron is relatively easily incorporated into diamond³; as boron acts as a charge acceptor, the resulting diamond is effectively hole-doped. Here we report the discovery of superconductivity in boron-doped diamond synthesized at high pressure (nearly 100,000 atmospheres) and temperature (2,500–2,800 K). Electrical resistivity, magnetic susceptibility, specific heat and field-dependent resistance measurements show that boron-doped diamond is a bulk, type-II superconductor below the superconducting transition temperature $T_c \approx 4$ K; superconductivity survives in a magnetic field up to $H_{c2}(0) \geq 3.5$ T. The discovery of superconductivity in diamond-structured carbon suggests that Si and Ge, which also form in the diamond structure, may similarly exhibit superconductivity under the appropriate conditions.

With their potential for electronic applications as microchip substrates, high-efficiency electron emitters, photodetectors and transistors, diamond and carrier-doped diamond have been studied extensively^{3–6}. The extremely short covalent bonds of carbon atoms in diamond give diamond many of its desirable properties, but also constrain geometrically which dopants can be incorporated and their concentration. Because of its small atomic radius compared to other potential dopants, boron is readily incorporated into the dense (1.763×10^{23} atoms cm^{-3}) diamond lattice. Boron dopes holes into a shallow acceptor level close to the top of the valence band that is separated from the conduction band of diamond by $E_g \approx 5.5$ eV. Electrical transport studies of B-doped diamond, including high-pressure synthesized crystals and CVD (chemical vapour deposition) films, find that low boron concentrations $n \approx 10^{17}–10^{19} \text{ cm}^{-3}$ give a semiconducting conductivity with an activation energy of ~ 0.35 eV (refs 7–11). Increasing the concentration to 10^{20} cm^{-3} gradually decreases the activation energy^{9,10}, and for $n \geq 10^{20} \text{ cm}^{-3}$, the electrical conductivity acquires metallic-like behaviour near room temperature^{8–11} that signals an insulator–metal transition near this concentration. A metallic-like conductivity has not been found, however, at low temperatures for any presently available B concentration, which has reached $(2–3) \times 10^{21} \text{ cm}^{-3}$ (refs 8–11).

We have studied B-doped diamond synthesized by reacting B_4C and graphitic carbon at pressure, 8–9 GPa, and temperature, 2,500–2,800 K, for ~ 5 s. Under these conditions, polycrystalline diamond aggregates 1–2 mm in size formed at the interface between graphite and B_4C . All the diamond aggregates had a metal-like lustre. Scanning electron microscopy (SEM) showed that the diamond grain size was a few micrometres (Fig. 1). The samples were

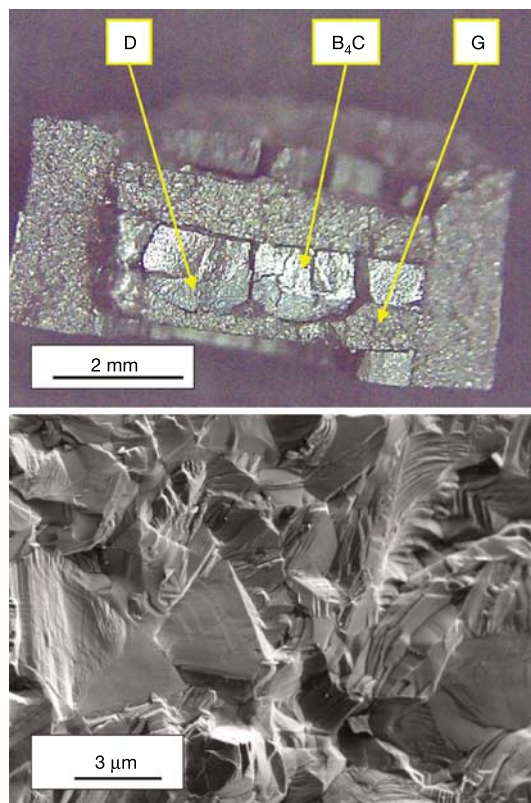


Figure 1 Optical and scanning electron microscopy images of the material. Top, central part of the high-pressure synthesis cell after subjecting graphite and B_4C to high-pressure, high-temperature conditions. D, diamond; G, graphite. Bottom, SEM image of B-doped diamond synthesized at high pressures and temperatures.

characterized by X-ray diffraction, micro-Raman spectroscopy, electrical resistivity, magnetic susceptibility and calorimetry measurements. Consistent with an expanded lattice deduced from the X-ray spectrum given in Fig. 2a, Raman spectra of our samples (Fig. 2b) are very similar to those observed for the most heavily boron-doped CVD diamond films¹¹. (From the weak zone-centre phonon line at $\sim 1,300\text{ cm}^{-1}$ and the general form of the spectra¹¹, our samples should have a carrier concentration $n \geq 2 \times 10^{21}\text{ cm}^{-3}$. For normal diamond, a zone-centre phonon line is observed at $1,332\text{ cm}^{-1}$.) Quantitative estimates of the boron content in our samples were made using NMR and inductively coupled mass spectrometry. Within experimental uncertainty, both methods yielded the same value of $2.8 \pm 0.5\%$ for the total B content. Inductively coupled mass spectrometry on a second sample also gave the same B content. Good agreement between these two techniques, one of which is volume sensitive (NMR) and the other of which probes only a surface layer of $\sim 1,000\text{ \AA}$, suggests that B atoms are distributed rather uniformly in the sample. Neither technique, however, is sensitive to the local environment of the B atoms. Assuming that all the B atoms are incorporated into the diamond lattice, the upper limit of the charge carrier concentration is $4.9 \times 10^{21}\text{ cm}^{-3}$, a value somewhat larger than the maximum previously reported. This estimate decreases to $4.6 \times 10^{21}\text{ cm}^{-3}$, assuming the sample contains 2% B_4C .

The temperature (T)-dependent resistivity (ρ) of one of these B-doped samples is plotted in Fig. 3a. Measurements of current-voltage characteristics at various temperatures confirmed an ohmic response. In the range 230–300 K, $\partial\rho/\partial T > 0$ as expected for a metal, but below 230 K the resistivity increases weakly with decreasing temperature. Near 4 K, $\rho(T)$ starts to fall rapidly and reaches an immeasurably small value below 2.3 K as shown in the inset of Fig. 3a. The resistive variation below 4 K is typical of an inhomogeneous superconductor, in this case with the inhomogeneity probably arising from a non-uniform distribution of boron in the diamond lattice. As also indicated in the inset of Fig. 3a, the application of hydrostatic pressure produces a linear decrease in

the resistive mid-point transition temperature T_c without any additional significant broadening of the transition. The relatively slow decrease of T_c with pressure P ($\partial T_c/\partial P = -6.42 \times 10^{-2}\text{ K GPa}^{-1}$, $\partial \ln T_c/\partial P = 2.79 \times 10^{-2}\text{ GPa}^{-1}$) contrasts with positive $\partial T_c/\partial P = 0.05\text{ K GPa}^{-1}$ reported¹² for elemental boron for $P > 180\text{ GPa}$, ruling out the possibility that the zero-resistance state in B-doped diamond is due to free boron. As expected for an inhomogeneous superconductor, applying a magnetic field slightly broadens the transition and shifts it to lower temperatures (Fig. 3b). The resistive mid-point T_c decreases at a rate $\partial H_{c2}/\partial T_c = -1.7\text{ T K}^{-1}$, from which we estimate the $T = 0$ upper critical field $H_{c2}(0) = 3.4\text{ T}$ using the standard relationship¹³ for a dirty type-II superconductor $H_{c2}(0) = -0.69 (dH_{c2}/dT|_{T_c}) T_c$. As is evident in the inset of Fig. 3b, this estimate is a lower limit because the onset of the resistive transition is still near 1.7 K in a field of 4 T. With $H_{c2}(0) = 3.4\text{ T}$ and the relation $\xi_{GL} = [\Phi_0/2\pi H_{c2}(0)]^{1/2}$ (where $\Phi_0 = hc/2e$ is the quantum of magnetic flux), the Ginzburg–Landau coherence length $\xi_{GL} = 100\text{ \AA}$.

Resistance measurements are unable to determine if superconductivity arises from the crystal bulk, surface, or filaments of zero-resistance material; however, magnetic susceptibility measurements allow more definitive conclusions. As shown in Fig. 4, there is a strong diamagnetic response in the a.c. magnetic susceptibility of these boron-doped diamond samples below 2.3 K where the resist-

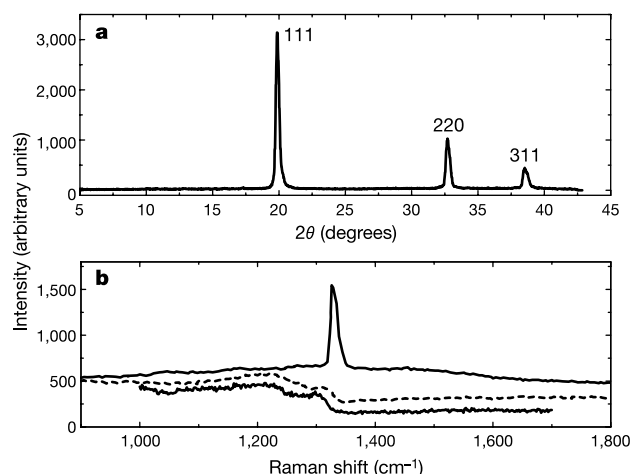


Figure 2 X-ray diffraction and Raman spectra of B-doped diamond. **a**, X-ray diffraction pattern of B-doped diamond. The strongest peaks of B_4C at $16\text{--}17^\circ$ and graphite near $12\text{--}13^\circ$ are absent in this pattern from diamond, but are visible in spectra (not shown) of material taken from the boundary between B_4C and diamond. We estimate that unreacted B_4C in the polycrystalline diamond does not exceed 1–2%. The diffraction pattern gives a cubic lattice parameter of $3.5755 \pm 0.0005\text{ \AA}$, which is larger than $3.5664\text{ \AA}$ for undoped diamond and within the range of lattice parameters $3.575\text{--}3.5767\text{ \AA}$ for maximally B-doped diamond^{24,25}. **b**, Raman spectra of diamond, synthesized in the system graphite– B_4C (lower curve) and CVD films¹¹ with boron concentrations of $3.4 \times 10^{17}\text{ cm}^{-3}$ (upper curve) and $1.5 \times 10^{21}\text{ cm}^{-3}$ (middle curve).

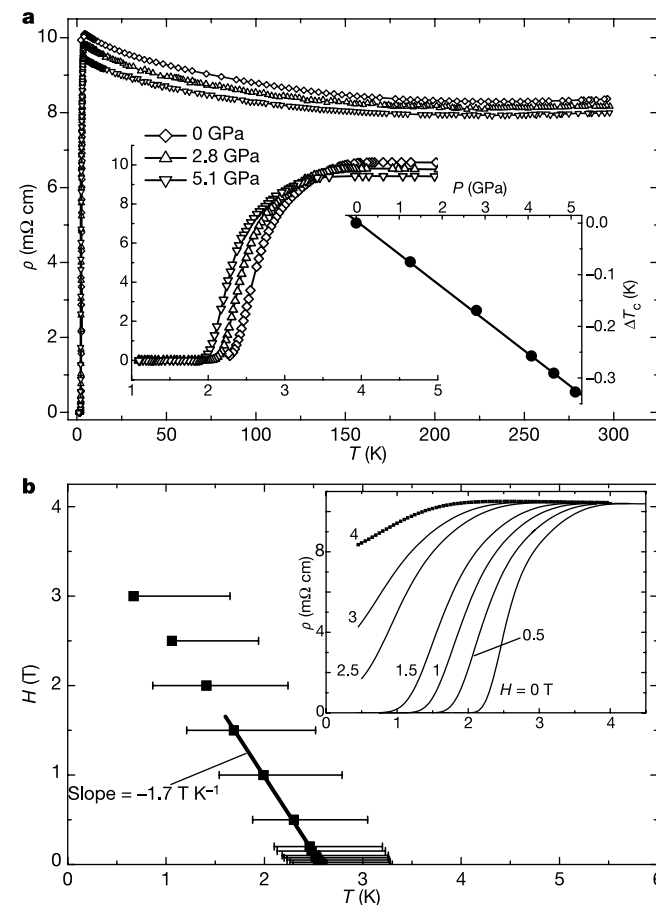


Figure 3 Electrical resistivity and upper critical field curves for B-doped diamond. **a**, Temperature dependence of the electrical resistivity of B-doped diamond at normal and representative high pressures. The insets show details of the resistivity behaviour below 5 K and the pressure-induced shift ΔT_c of the midpoint of the resistive transition. **b**, Temperature dependence of the upper critical field for B-doped diamond. The resistive mid-point is used to define H_{c2} . The inset shows the evolution of the resistivity near T_c at different magnetic fields 0–4 T.

ance drops to zero. At 1.1 K, the diamagnetic response for various B-doped samples ranged from 25% to 60% of the value measured for a superconducting indium sample in the same coil system. An absolute value of magnetic susceptibility for two separate samples, measured by SQUID (superconducting quantum interference device) magnetometry, agrees with the above estimates (Fig. 4), and, further, finds a diamagnetic response on cooling in a d.c. field of 5 Oe. These results indicate that the superconductivity is not filamentary. Moreover, magnetization measurements versus field (inset of Fig. 4) find the magnetic hysteresis expected in a bulk superconductor with flux pinning.

A specific heat anomaly at T_c would provide the most definitive evidence that superconductivity develops in the bulk of B-doped diamond. Assuming that the effective mass of charge carriers is the free electron mass m_e , we estimate the electronic contribution to the specific heat as $C = \gamma T$. $\gamma = 7.6 \times 10^{-4} \text{ J mol}^{-1} \text{ K}^{-2}$ from the free electron relationships $k_F = (3\pi^2 n)^{1/3}$ and $\gamma = \pi^2 n k_B^2 m_e / \hbar^2 k_F^2$, where k_F is the Fermi momentum, n is the carrier density, and k_B is Boltzmann's constant. For a weak coupled BCS superconductor, the jump in specific heat C at T_c is $\Delta C / \gamma T_c = 1.43$, which implies $\Delta C = 2.5 \times 10^{-3} \text{ J mol}^{-1} \text{ K}^{-1}$ and for our largest 2-mg sample, then, $\Delta C = 4.2 \times 10^{-7} \text{ J K}^{-1}$. This is a very small value, essentially equal to the noise level of a commercial Quantum Design Physical Property Measurement System. We have measured the specific heat of this 2-mg sample in such a device and find a barely discernable anomaly near 2.3 K.

For phonon-mediated superconductivity, T_c is given approximately by $T_c = (\theta_D / 1.45) \exp(-1.04(1 + \lambda) / \lambda)$, where θ_D is the Debye temperature and λ is a measure of electron-phonon coupling¹⁴. The high Debye temperature of diamond ($\theta_D = 1,860 \text{ K}$) suggests that carrier doping the diamond lattice could produce a relatively high T_c . Using this value of θ_D and $T_c = 2.3 \text{ K}$, we estimate $\lambda = 0.2$, indicative of weak electron-phonon coupling. This somewhat small value of λ can be understood qualitatively from the relation $\lambda = N(0) \langle I^2 \rangle / (M \langle \omega^2 \rangle)$, where $N(0)$ is the density of electronic states at the Fermi energy, $\langle I^2 \rangle$ is the average square of the electron-phonon matrix element, M is the ionic mass and $\langle \omega^2 \rangle$ is a characteristic phonon frequency averaged over the phonon spectrum¹⁴. Although the ionic mass is light, the strong covalent bonds in diamond and low concentration of charge carriers imply a small $N(0)$ and large $\langle \omega^2 \rangle$.

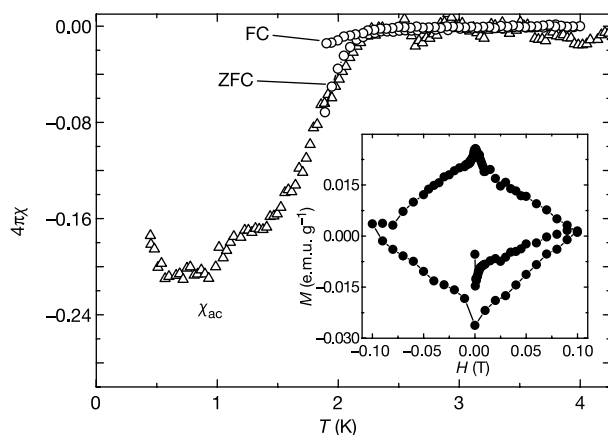


Figure 4 Magnetic susceptibility of B-doped diamond near T_c . The a.c. data (triangles) were normalized to the d.c. data (circles) at 1.8 K. ZFC, zero-field cooled; FC, field cooled. No diamagnetic signal was found in B_4C samples prepared at the same pressure and temperature conditions used for the diamond synthesis. The inset plots magnetization versus field at 1.8 K. From the magnitude of the hysteresis at $H = 0$, a simple estimate based on Bean's critical state model²⁶ gives a critical current density of $\sim 1.45 \times 10^6 \text{ A m}^{-2}$.

Superconductivity in doped semiconductors is an intriguing issue for theory^{15–19}, but experimentally only a few superconducting semiconductors have been discovered. Bulk superconductivity with $T_c \approx 0.1\text{--}0.5 \text{ K}$ was found²⁰ in self-doped GeTe and SnTe at carrier densities $n \approx 10^{21} \text{ cm}^{-3}$ and in doped SrTiO_3 ($T_c \approx 0.05\text{--}0.25 \text{ K}$ at $n \approx 10^{19}\text{--}10^{21} \text{ cm}^{-3}$). There have been no reports until now of superconductivity in group-IV semiconductors with the diamond structure—silicon, germanium and their alloys—though it was predicted¹⁶ to exist at very low temperatures. The discovery of bulk superconductivity in B-doped diamond should stimulate a fresh look at this old problem, which is posed more straightforwardly in B-doped diamond than in the more complex Si- or Ge-based clathrates, which are the first superconductors^{21,22} built from a network of tetrahedral covalent bonds having bond lengths comparable to those of cubic diamond. Few, if any, measurements of B-doped diamond have been made below 4 K (ref. 8), so that superconductivity may have been missed in samples prepared by others with lower B concentrations. A systematic study seems worthwhile. Our results also suggest that homogeneously doping diamond with B concentrations of $(4\text{--}5) \times 10^{21} \text{ cm}^{-3}$ should produce sharp transitions near 4 K and upper critical fields above 4 T. □

Methods

Sample preparation and analysis

Starting materials, graphite discs containing $\leq 500 \text{ p.p.m.}$ impurities and B_4C powder containing less than 3,000 p.p.m. impurities (mainly $\text{Fe} \leq 2,000 \text{ p.p.m.}$, and $\text{Si} \leq 1,000 \text{ p.p.m.}$), were placed in a graphite heater assembly. Impurity content in the starting materials was determined by X-ray fluorescence and spectroscopic methods, both showing identical results. Starting materials were annealed in vacuum to remove moisture. The heater was insulated by boron nitride sleeves from the gasket material (lithographic stone, mainly $\text{CaCO}_3 + \text{SiO}_2$) and opposed tungsten carbide anvils (with 6% Co) of a high-pressure toroid apparatus (ref. 23). Molybdenum disks separated cylindrical parts of the graphite heater and anvils. Impurity content in the synthesized diamond was determined by X-ray fluorescence, with special attention to Ca, Co, Fe and Si, which could arise from the gasket, anvils and B_4C . Total content of these elements does not exceed 500 p.p.m. The boron content in our diamond samples was estimated in two ways. NMR spectra were measured at room temperature and peaks corresponding to ^{13}C and ^{11}B isotopes were observed. The ratio of boron/carbon atoms in the sample was calculated on the basis of the peak areas and the known natural abundance of $^{11}\text{B}/^{13}\text{B}$ and $^{13}\text{C}/^{12}\text{C}$ isotopes. The second method used a high-resolution inductively coupled mass spectrometer (ThermoFinnigan-MAT 'Element 2') equipped with an ultraviolet laser ablation microprobe. The samples also were characterized by X-ray diffraction and micro-Raman spectroscopy. X-ray diffraction studies were made with a BRUKER-AXS diffractometer and monochromatized Mo radiation. Micro-Raman spectroscopy was performed using a U-1000 spectrometer. The laser exciting radiation (488 nm) was focused to a spot $\sim 20 \mu\text{m}$ in diameter. The Raman spectra were recorded in backscattering geometry with a resolution better than 1 cm^{-1} .

Electric and magnetic measurements

Resistivity measurements were made in a standard four-terminal configuration using an LR700 a.c. resistance bridge. Silver epoxy attached 25- μm Pt wires to the samples. Measurements in the range 1.1–300 K were performed in a ^4He cryostat. Resistivity measurements at hydrostatic pressures to 5.3 GPa were performed in a miniature clamped toroidal cell. A diamond sample and lead manometer were placed inside a small Teflon capsule filled with liquid. Pressure was applied at room temperature and then the clamped cell was cooled slowly to 1.1 K. Measurements of the resistivity in the range 0.5–4 K and in magnetic fields to 4 T were performed in a Quantum Design Physical Properties Measurement System. Measurements of a.c. magnetic susceptibility were made in two apparatuses, one to ^4He and the other to ^3He temperatures, by measuring the mutual inductance of a miniature coil with an LR700 bridge. The coil response was calibrated by measuring the susceptibility of a piece of superconducting indium, having a shape and size similar to that of diamond samples, or by normalizing the data to d.c. susceptibility results obtained for $T \geq 1.8 \text{ K}$ in a Quantum Design Magnetic Properties Measurement System with an applied magnetic field of 5 Oe.

Received 23 January; accepted 1 March 2004; doi:10.1038/nature02449.

- Field, J. E. *The Properties of Natural and Synthetic Diamond* (Academic, New York, 1992).
- May, P. W. Diamond thin films: a 21st-century material. *Phil. Trans. R. Soc. Lond. A* **358**, 473–495 (2000).
- Kalish, R. The search for donors in diamond. *Diamond Relat. Mater.* **10**, 1749–1755 (2001).
- Geis, M. W., Twitchell, J. C., Macaulay, J. & Okano, K. Electron field emission from diamond and other carbon materials after H_2 , O_2 and Cs treatment. *Appl. Phys. Lett.* **67**, 1328–1330 (1995).
- Teukam, Z. et al. Shallow donors with high n-type electrical conductivity in homoepitaxial deuterated boron-doped diamond layers. *Nature Mater.* **2**, 482–486 (2003).
- Garrido, J. A. et al. Fabrication of in-plane gate transistors on hydrogenated diamond surfaces. *Appl. Phys. Lett.* **82**, 988–990 (2003).

7. Williams, A. W. S., Lightowers, E. C. & Collins, A. T. Impurity conduction in synthetic semiconducting diamond. *J. Phys. C* **3**, 1727–1735 (1970).
8. Vishnevskii, A. S., Gontar', A. G., Torishnii, V. I. & Shul'zhenko, A. A. Electrical conductivity of heavily doped p-type diamond. *Sov. Phys. Semicond.* **15**, 659–661 (1981).
9. Werner, M. *et al.* Charge transport in heavily B-doped, polycrystalline diamond films. *Appl. Phys. Lett.* **64**, 595–597 (1994).
10. Borst, T. H. & Weis, O. Boron-doped homoepitaxial diamond layers: fabrication, characterization and electronic applications. *Phys. Status Solidi A* **154**, 423–444 (1996).
11. Zhang, R. J., Lee, S. T. & Lam, Y. W. Characterization of heavily boron-doped diamond films. *Diamond Relat. Mater.* **5**, 1288–1294 (1996).
12. Eremets, M. I., Struzhkin, V. V., Mao, H.-K. & Hemley, R. J. Superconductivity in boron. *Science* **293**, 272–274 (2001).
13. Werthamer, N. R., Helfand, E. & Hohenberg, P. C. Temperature and purity dependence of the superconducting critical field H_{c2} . III. Electron spin and spin-orbit effects. *Phys. Rev.* **147**, 295–302 (1966).
14. McMillan, W. L. Transition temperature of strongly coupled superconductors. *Phys. Rev.* **167**, 331–344 (1968).
15. Gurevich, V. L., Larkin, A. I. & Firsov, Yu. A. On the possibility of superconductivity in semiconductors. *Sov. Phys. Solid State* **4**, 131–135 (1962).
16. Cohen, M. L. Superconductivity in many-valley semiconductors and in semimetals. *Phys. Rev.* **134**, A511–A521 (1964).
17. Cohen, M. L. in *Superconductivity* (ed. Parks, R. D.) Vol. 1, 615–644 (Marcel Dekker, New York, 1969).
18. Kohmoto, M. & Takada, Y. Superconductivity from an insulator. *J. Phys. Soc. Jpn* **59**, 1541–1544 (1990).
19. Nozières, P. & Pistoiesi, F. From semiconductors to superconductors: a simple model for pseudogaps. *Eur. Phys. J. B* **10**, 649–662 (1999).
20. Hulm, J. K., Ashkin, M., Deis, D. W. & Jones, C. K. in *Progress in Low Temperature Physics* (ed. Gorter, C. J.) Vol. VI, 205–242 (North-Holland, Amsterdam, 1970).
21. Kawaji, H., Horie, H., Yamanaka, S. & Ishikawa, M. Superconductivity in the silicon clathrate compound $(\text{Na,Ba})_x\text{Si}_{46}$. *Phys. Rev. Lett.* **74**, 1427–1429 (1995).
22. Grosche, F. M. *et al.* Superconductivity in the filled cage compounds $\text{Ba}_6\text{Ge}_{25}$ and $\text{Ba}_4\text{Na}_2\text{Ge}_{25}$. *Phys. Rev. Lett.* **87**, 247003 (2001).
23. Khvostantsev, L. G., Vereshchagin, L. F. & Novikov, A. P. Device of toroid type for high pressure generation. *High Temp. High Press.* **9**, 637–639 (1977).
24. Voronov, O. A. & Rakhmanina, A. V. Parameter of the cubic cell of diamond doped with boron. *Inorg. Mater.* **29**, 707–710 (1993).
25. Brunet, F., Deneuville, A., Germa, P., Pernet, M. & Gheeraert, E. Variation of the cell parameter of polycrystalline boron doped diamond films. *J. Appl. Phys.* **81**, 1120–1125 (1997).
26. Bean, C. P. Magnetization of hard superconductors. *Phys. Rev. Lett.* **8**, 250–253 (1962).

Acknowledgements We thank D. Wayne for mass spectrometry measurements of the B content of our samples, A. Presz for SEM images and S. Gierlotka for help in sample analysis. This work was supported by the Russian Foundation for Basic Research and by the Strongly Correlated Electrons Program of the Department of Physical Sciences, Russian Academy of Sciences. Work at Los Alamos was performed under the auspices of the US DOE.

Authors' contributions Boron-doped diamond samples were synthesized by E.A.E., and their physical properties measured by V.A.S., E.D.B., N.N.M., N.J.C., J.D.T. and S.M.S.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to V.A.S. (sidorov@hppi.troitsk.ru)

Intermediate-depth earthquake faulting by dehydration embrittlement with negative volume change

Haemyeong Jung¹, Harry W. Green II^{1,2} & Larissa F. Dobrzhinetskaya²

¹Institute of Geophysics and Planetary Physics, ²Department of Earth Sciences, University of California, Riverside, California 92521, USA

Earthquakes are observed to occur in subduction zones to depths of approximately 680 km, even though unassisted brittle failure is inhibited at depths greater than about 50 km, owing to the high pressures and temperatures^{1–3}. It is thought that such earthquakes (particularly those at intermediate depths of 50–300 km)

may instead be triggered by embrittlement accompanying dehydration of hydrous minerals, principally serpentine^{1–3}. A problem with failure by serpentine dehydration is that the volume change accompanying dehydration becomes negative at pressures of 2–4 GPa (60–120 km depth), above which brittle fracture mechanics predicts that the instability should be quenched^{4,5}. Here we show that dehydration of antigorite serpentine under stress results in faults delineated by ultrafine-grained solid reaction products formed during dehydration. This phenomenon was observed under all conditions tested (pressures of 1–6 GPa; temperatures of 650–820 °C), independent of the sign of the volume change of reaction. Although this result contradicts expectations from fracture mechanics, it can be explained by separation of fluid from solid residue before and during faulting, a hypothesis supported by our observations. These observations confirm that dehydration embrittlement is a viable mechanism for nucleating earthquakes independent of depth, as long as there are hydrous minerals breaking down under a differential stress.

A popular hypothesis for overcoming brittle fracture inhibition, especially for earthquakes at intermediate depths (<300 km), is assistance of brittle fracture by generation of a free fluid as a result of dehydration of serpentine or other hydrous minerals^{1–3,6–11}. The phenomenon of dehydration embrittlement was discovered almost 40 years ago¹ but has been studied only sporadically since that time^{4–8}. In particular, this phenomenon has not been addressed by controlled deformation experiments at pressures greater than 700 MPa, equivalent to only ~20 km depth in Earth. Studies of acoustic emission at elevated pressures^{5,8}, however, have implications for deeper earthquakes. The latter studies, although lacking control of differential stress, strain, or strain rate, recorded acoustic emissions at much higher pressures and inferred that faulting had occurred.

The fracture mechanics explanation of how dehydration embrittlement can enable brittle shear failure at elevated pressure is based upon production of a pore pressure as a consequence of a positive volume change, ΔV , of the dehydration reaction and consequent decrease in the effective pressure on existing or potential planes of weakness. Thus, as conventionally understood, the theory predicts that if the ΔV of the reaction were to become negative, failure would become more difficult and the shearing instability would vanish^{4,5}. Because hydrous fluid is much more compressible than solid silicates, the total ΔV of dehydration of the common hydrous phases of primary interest (for example, serpentine and chlorite) is progressively reduced as pressure increases and becomes negative at pressures of 2–4 GPa, equivalent to pressures of 60–120 km in Earth. One consequence of this prediction is that earthquakes should not be possible by dehydration embrittlement at greater depths⁵. However, dehydration embrittlement is the only earthquake nucleation mechanism known to be viable for depths less than 300 km (ref. 2). Thus, it is important to determine whether this mechanism can function under conditions where ΔV is negative.

We report here the results of deformation experiments at pressures of 1–6 GPa and temperatures of 550–820 °C using an antigorite serpentine from Val Malenco, Italy, for which the phase diagram has been measured¹². Figure 1a shows the experimental conditions investigated; the slope of the high-temperature limit of antigorite stability (Fig. 1a) is negative above ~2.2 GPa, reflecting negative ΔV of reaction above that pressure.

Figure 1b–f shows microstructures of the starting material and results of annealing without deformation outside antigorite stability. The layering, strong foliation and proportions of antigorite and relict olivine shown are typical of our starting material. Figure 1f shows breakdown of antigorite along boundaries with relict olivine and healed cracks of several orientations outlined by fluid inclusions; such inclusion trails are rare in the starting material,

hence clearly were produced during the experiment. Similar healed cracks are abundant and strongly oriented in deformed specimens (see below).

One specimen was deformed within the stability field of antigorite (triangle in Fig. 1a). This specimen was extremely strong ($\sigma_1 - \sigma_3 = 1.5$ GPa) and failed by conventional brittle fracture; no sign of dehydration products was found within the specimen after deformation (σ_1 is maximum compressive stress; σ_3 is minimum compressive stress). All other specimens deformed in the Griggs apparatus failed at significantly lower stresses (for example, 0.8–0.9 GPa at $\sigma_3 = 1$ GPa), exhibiting multiple faults. Except in complicated multi-fault regions, faults are aligned ~ 25 – 40° to the maximum applied stress, σ_1 . Figure 2a–c shows fault displacements demonstrated by offsets of specimen surfaces, regions of relict olivine, and across films of Ni incorporated as described in Methods. Figure 2d–f shows details of the breakdown of antigorite adjacent to relict olivine where progression of the reaction can be

seen much more clearly than in the antigorite matrix. Many grain boundaries are open and show cusps along them, reflecting dissolution of olivine by the fluid released during antigorite decomposition. Open Mode I cracks parallel to σ_1 (top–bottom in photographs) are abundant, as are healed cracks outlined by fluid inclusion trails. Open cracks display cusps with smaller cracks, commonly healed, emanating from them.

Faults all contain ultrafine-grained aggregates (≤ 300 nm) of solid dehydration products that consist primarily of olivine, as shown by energy-dispersive spectroscopy (EDS) analysis and bright contrast. Although the ΔV of reaction is strongly positive at 1.0 GPa, moderately positive at 1.7 GPa and distinctly negative at 3.3 and 6 GPa, all specimens showed the same ‘wispy’ pattern of solid reaction products preserved within and along the faults.

Figure 3a (pressure $P = 1.7$ GPa) shows a fault that branches near the centre of the image. The branch to the right ends near the top of the image (arrowheads), with the wispy dehydration products changing orientation to approximately normal to macroscopic σ_1 (top–bottom), defining surfaces akin to stylolites¹³ or anticracks¹⁴ (surfaces across which contraction has occurred and volume has been lost). The right-lateral shear on this fault system is such that the triangular region near the branch (arrow) is in compression and displays a high density of dehydration product lenses. It is well known that faults at Earth’s surface commonly show *en echelon* fault segments, with ‘pull-aparts’ between them¹⁵. The inset of Fig. 3a shows an analogous feature that is very common in our experiments in which, rather than a ‘pull-apart’, a ‘push-together’ is formed between *en echelon* fault segments arranged such that the region between their overlapping tips has experienced contraction rather than extension. The region of the ‘push-together’ has an abundance of dehydration products defining anticrack lenses that have accommodated shortening in response to the enhanced compression. Figure 3b ($P = 6$ GPa) shows a series of such ‘push-togethers’ (also shown diagrammatically in Fig. 3c for clarity), again showing

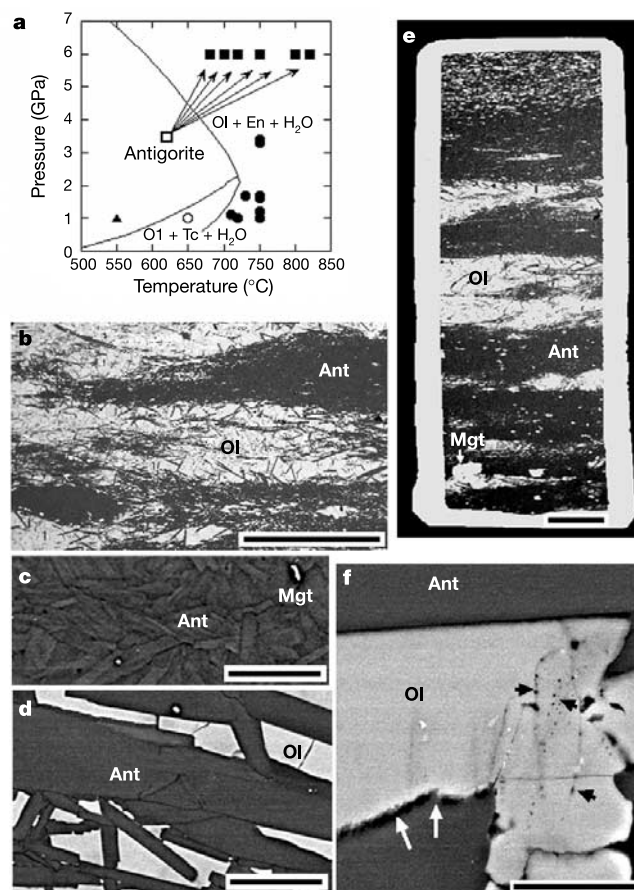


Figure 1 Experimental conditions and starting material for deformation experiments. **a**, Phase diagram for antigorite¹² and experimental conditions. Ol, olivine; En, enstatite; Tc, talc. Open symbols, hydrostatic experiments. Filled circles, deformation experiments in modified Griggs apparatus (strain rate $2 \times 10^{-4} \text{ s}^{-1}$). Filled squares, experiments in Walker-type multianvil apparatus; arrows indicate P/T paths during rapid pumping, yielding strain rates approximating those in the Griggs apparatus. All deformation experiments displayed faults and evidence of antigorite breakdown except the experiment within antigorite stability (filled triangle). **b–d**, Microstructure of starting material: antigorite (Ant), relict olivine (Ol), and minor magnetite (Mgt). **e, f**, Specimen pressurized to 1.0 GPa and annealed for 1 h at 650 °C (open circle in **a**). Pt capsule (white) and specimen are undeformed. **f**, Minor antigorite breakdown (arrows) yields fluid inclusion trails (arrowheads) on healed cracks of several orientations. Scale bars: 1 mm (**b, e**); 10 μm (**c, d, f**).

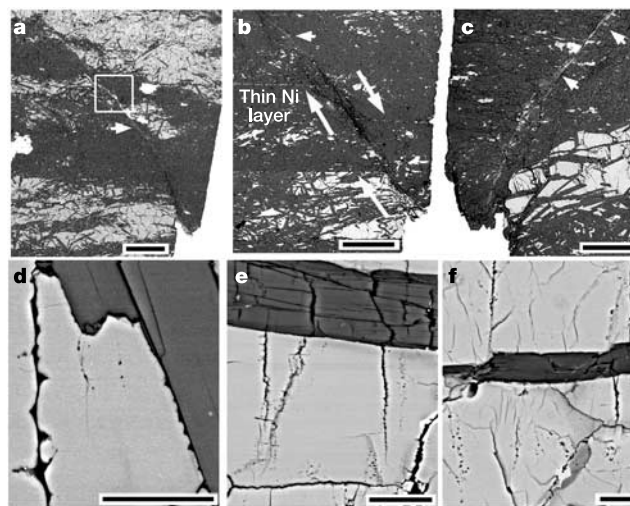


Figure 2 Microstructures of faulted specimens. **a**, Fault offsetting layers richer and poorer in relict olivine (1.7 GPa, 750 °C; $\Delta V > 0$); square shows location of Fig. 3a. **b**, Fault displacing Ni films (arrows) (1.0 GPa, 720 °C; $\Delta V > 0$). **c**, Similar to **a** but 3.3 GPa, 750 °C; $\Delta V < 0$. Dehydration products along faults in **a–c** indicated by arrowheads. **d–f**, Open Mode I cracks (σ_1 is top–bottom in image) with multiple cusps showing dissolution of olivine and fluid inclusion trails outlining healed Mode I cracks. Minor irregular cracking during decompression after experiment also visible. **d**, 1.7 GPa, 750 °C; $\Delta V > 0$. **e**, 1.7 GPa, 730 °C; $\Delta V > 0$. **f**, 6 GPa, 750 °C; $\Delta V < 0$. Scale bars: 200 μm (**a–c**); 10 μm (**d**); 20 μm (**e, f**).

that the dehydration products decorate surfaces of maximum compression. Figure 3d ($P = 1.0$ GPa) shows another fault trace that displays the same features as described in Fig. 3a–c. In this case, branching of the fault zone is seen at the top left corner of the image, with compression (anticrack) lenses in the squeezed wedge at the branch point (star), and several ‘push-togethers’, marked by arrowheads, are strung out along the fault trace. These panels also show that the extent of reaction outside fault zones is distinctly less than within them, suggesting that the fault zones have self-organized and propagated from points where the reaction initiated.

In summary, dehydration of antigorite under stress at pressures of 1–6 GPa leads to faulting under all conditions tested. Moreover, we document the presence of abundant healed cracks marked by fluid inclusion trails on surfaces of least normal stress (parallel to σ_1) and lenses of solid reaction products located preferentially on surfaces of highest normal stress (normal to σ_1), showing that the fluid and solid products of antigorite breakdown become separated during faulting. Thus, free fluid of density lower than antigorite was

available at all pressures to generate abundant Mode I cracks and lead to shear failure². Faulting apparently in conflict with fracture mechanics theory is explained by separation of fluid (lower density than antigorite—positive ΔV) and solid reaction products (higher density than antigorite—negative ΔV) during deformation. Each component is concentrated on appropriate surfaces to help accommodate the local strain field.

Returning to consideration of Fig. 3, we note that during growth of a fault (or growth of a slip patch on an existing fault), at any point in that growth the stress distribution resulting from its displacement field is as sketched in Fig. 4a. As the slipping patch grows, there is continuing renewal of such stressed regions, with the compressive zones sketched in Fig. 4b, c. Thus, such a fault patch evolving in a system in which an on-going reaction produces a fluid/solid slurry is ideally situated to focus the reaction in local regions of enhanced compression, leading to concentration of the denser solids on planes of high normal stress as microanticracks and expelling the less-dense fluid into the fault zone where it reduces friction and lubricates the fault surface. The result of such a scenario would be

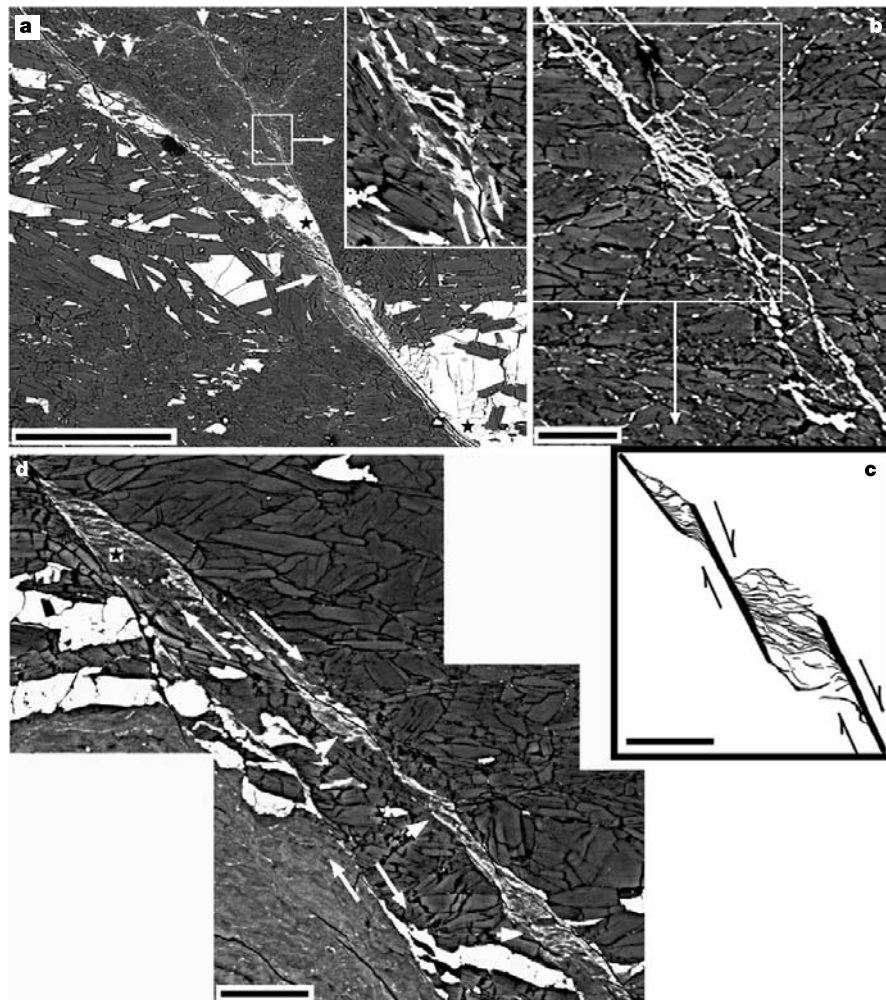


Figure 3 Fault microstructures. **a**, Detail of Fig. 2a. Fault shows minor branch that dies out near top of image, outlined by ‘wisps’ of solid reaction products at high angles to σ_1 (arrowheads). Convergence near branch point (arrow) is under compression and exhibits a concentration of ‘wisps’. Stars show fault displacement of a severed magnetite blob. Inset shows multiple anticrack ‘wisps’ in ‘push-together’ between overlapping ends of *en echelon* microfaults (1.7 GPa; $\Delta V > 0$). **b**, Topologically identical series of ‘push-togethers’ between *en echelon* fault segments at 6 GPa ($\Delta V < 0$). **c**, Schematic

drawing of part of fault zone in **b**; heavier lines are fault segments and lighter lines are traces of the solid reaction products. **d**, Similar microstructures at 1.0 GPa ($\Delta V > 0$); ‘push-togethers’ (arrowheads) and a converging zone (star) are again decorated by solid reaction products oriented normal to the orientation of the local σ_1 (top–bottom). Open cracks in the fault zones are decompression features. Scale bars: 100 μm (**a**); 20 μm (**b–d**).

microstructures like those shown in Fig. 3. Remarkably similar structures (except showing 'pull-aparts') have been shown in small natural faults^{13,14} (Fig. 4d), and 'push-togethers' have been seen previously between the tips of *en echelon* faults produced by transformation-induced faulting in Mg_2GeO_4 during the olivine-spinel transformation¹⁶.

These results show that dehydration embrittlement as a mechanism for initiation of earthquakes is not restricted to conditions where the total volume change of reaction is positive ($\Delta V > 0$). It is clear that in our experiments the self-organization of the shear failure process includes separation of fluid from solid reaction products, thereby extending the potential operation of pore-pressure-induced faulting to all conditions where the density of the fluid

generated is less than that of the solid phase from which the fluid is derived. That would surely include all conditions obtaining in Earth's upper mantle except perhaps at simultaneously high pressures and temperatures where solubility of rock components in fluid becomes very high¹⁷, potentially yielding melts more dense than the rock from which they are derived¹⁸.

Arguments that dehydration embrittlement can explain all intermediate-depth earthquakes (and perhaps deep earthquakes as well) have become common in recent years, both from the standpoint of seismology¹⁹ and petrology^{9,11,20}. Our results are supportive of this concept if the requisite dehydration reaction is occurring. On the other hand, serious questions have been raised about the ability of this mechanism to explain the lower plane of earthquakes in double seismic zones because the presence of hydrous phases would be required to depths of 40 km or more immediately beneath oceanic trenches. One potential resolution of this quandary is that deep penetration of H_2O perhaps occurs on major faults of the oceanic lithosphere, leading to hydrous alteration along them to the required depths^{19,21}. However, a recent study of the Vanuatu subduction zone²² shows that new faults are generated during subduction, an observation that would require more extensive hydrous alteration than solely along pre-existing faults. Reference 20 suggests up-dip flow of fluid along the dehydration reaction boundary, a mechanism that might be plausible, but to produce earthquakes directly below the trench, the fluid would have to continue to flow beyond the point where any subduction-related hydration could have reached that depth. The growing sophistication of seismic studies of subduction zones holds great promise for measuring the distribution of hydrous phases²³.

Lastly, recent laboratory experiments have demonstrated that generation of less than 1% fluid, if appropriately distributed, can be sufficient to enable faulting²⁴. As a consequence, it is possible that dehydration of even minor hydrous phases such as clinohumite in peridotite or phengite in eclogite could trigger earthquakes as deep as 400 km (ref. 25). Moreover, discovery of seismic reflectors that move²⁶ in fault zones of the continental crust clearly demonstrates fluid saturation in such fault zones. Therefore, hydrous alteration of crustal minerals must be widespread at seismogenic depths in continents, and dehydration of clays or other hydrous minerals just before (or in the early stages of) crustal faulting could also contribute significantly to initiation or enhancement of shallow earthquakes. □

Methods

Several studies of antigorite stability have yielded varying results that are attributed to variations in structure and Al content²⁷. For the results presented here, that uncertainty is not a problem for two reasons: (1) our results are dependent only on the running of the dehydration reaction, not the specific location of phase boundaries; (2) our starting material is the same as that used in ref. 12, hence we know the positions of the phase boundaries (Fig. 1a).

Experiments at pressures ≤ 3.4 GPa were conducted in a modified Griggs apparatus with sample assemblies illustrated previously^{24,28}. Samples were right circular cylinders, 3.1 mm diameter and 8.4 mm long, encapsulated in Pt with oxygen fugacity buffered at Ni/NiO. CsCl, an extremely weak solid, was used as pressure medium to minimize friction of the apparatus. In order to document small fault offsets, some experiments were composed of stacks of 20–25 disks with Ni evaporated on one surface of each disk. Pressure was increased to 40 MPa at room temperature with further pressurization at 300 °C. The temperature was then increased to the desired temperature, annealed for 0–60 min, and deformed at a strain rate of $2 \times 10^{-4} \text{ s}^{-1}$. Samples were quenched to room temperature in seconds following deformation. Each experiment contained thermocouples at both the top and bottom of the specimen; the temperature difference between them was generally less than $\sim 10^\circ\text{C}$. Experiments at pressures of 3.5–6 GPa were conducted using a Walker-type multi-anvil device with a single thermocouple in the centre of the specimen. The multi-anvil is designed to generate hydrostatic stresses. However, with Al_2O_3 pistons at the ends of fully-dense specimens, rapid pumping can generate the necessary differential stress to produce faulting²⁹. Therefore, samples were initially pressurized cold to 3.5 GPa and temperature raised to 620 °C followed by rapid simultaneous increase in pressure and temperature along the arrows shown in Fig. 1a over 5–9 min. No measurement of stress is possible in specimens deformed in the multi-anvil but, as shown in Figs 2 and 3, microstructures of faulted specimens were indistinguishable from those in the Griggs rig. Some specimens were made up of stacked disks of starting material on which Ni

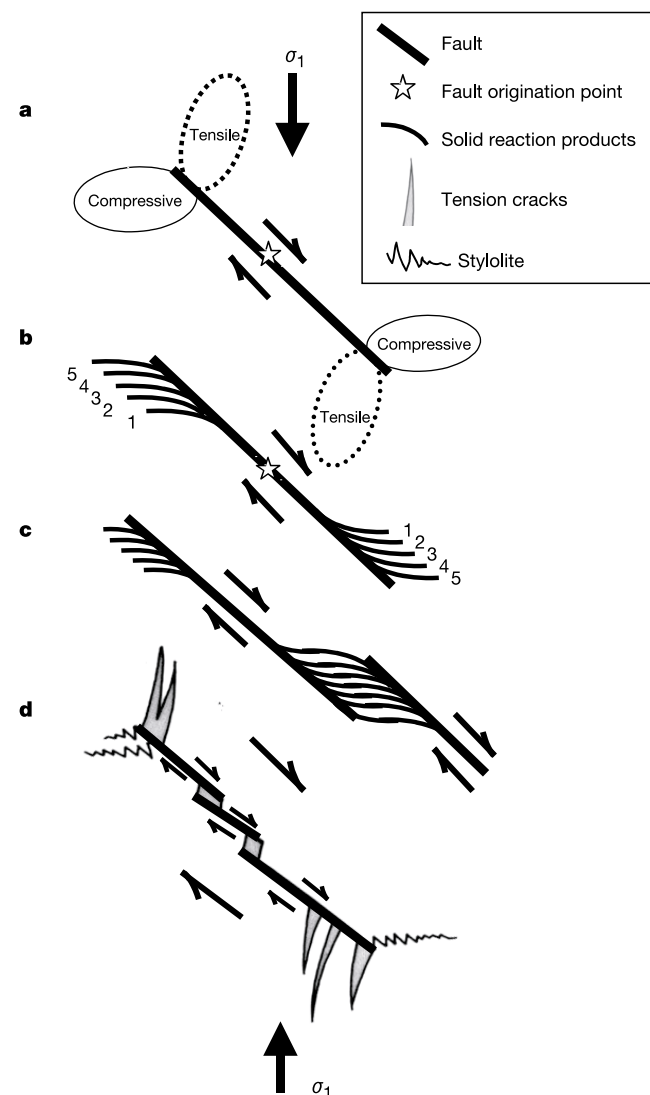


Figure 4 Hypothetical evolution of observed fault microstructure. **a**, Stress distribution at tips of growing fault. **b**, Potential microstructure developed along a newly formed microfault in which solid components ($\Delta V < 0$) are preserved periodically (sequential positions marked 1–5) and fluid components ($\Delta V > 0$) are not preserved (perhaps pumped into the evolving fault zone). Note that solid reaction products are oriented normal to local σ_1 . **c**, *En echelon* array of microfaults with characteristics illustrated in **b**. **d**, Sketched natural fault in limestone (modified after Fig. 3 of ref. 13) showing both tension cracks (filled with calcite) and compression stylolites. Note that *en echelon* stepping is the reverse sense from our high-pressure faults, yielding 'pull-aparts' (also filled with calcite).

films were evaporated on the disk surfaces, yielding well-defined internal markers of fault offsets. All deformed specimens were observed by optical and scanning electron microscopy (SEM); all half-tone figures are back-scattered electron images taken at 20 kV acceleration voltage with a Philips XL-30 SEM with a field emission gun.

Received 17 November 2003; accepted 17 February 2004; doi:10.1038/nature02412.

1. Raleigh, C. B. & Paterson, M. S. Experimental deformation of serpentinite and its tectonic implications. *J. Geophys. Res.* **70**, 3965–3985 (1965).
2. Green, H. W. & Houston, H. The mechanics of deep earthquakes. *Annu. Rev. Earth Planet. Sci.* **23**, 169–213 (1995).
3. Kirby, S. Intralab earthquakes and phase changes in subducting lithosphere. *Rev. Geophys.* **33**(suppl.), 287–297 (1995).
4. Wong, T.-F., Ko, S. C. & Olgaard, D. L. Generation and maintenance of pore pressure excess in a dehydrating system, 2. Theoretical analysis. *J. Geophys. Res.* **102**, 841–852 (1997).
5. Dobson, D. P., Meredith, P. G. & Boon, S. A. Simulation of subduction zone seismicity by dehydration of serpentine. *Science* **298**, 1407–1410 (2002).
6. Murrell, S. A. F. & Ismail, I. A. H. The effect of decomposition of hydrous minerals on the mechanical properties of rocks at high pressures and temperatures. *Tectonophysics* **31**, 207–258 (1976).
7. Rutter, E. H. & Brodie, K. H. Experimental “syntectonic” dehydration of serpentinite under conditions of controlled pore water pressure. *J. Geophys. Res.* **93**, 4907–4932 (1988).
8. Meade, C. & Jeanloz, R. Deep-focus earthquakes and recycling of water into the earth’s mantle. *Science* **252**, 68–72 (1991).
9. Peacock, S. Are the lower planes of double seismic zones caused by serpentine dehydration in subducting oceanic mantle? *Geology* **29**, 299–302 (2001).
10. Seno, T., Zhao, D., Kobayashi, Y. & Nakamura, M. Dehydration of serpentinized slab mantle: seismic evidence from southwest Japan. *Earth Planets Space* **53**, 861–871 (2001).
11. Omori, S., Kamiya, S., Maruyama, S. & Zhao, D. Morphology of the intraslab seismic zone and devolatilization phase equilibria of the subducting slab peridotite. *Bull. Earthq. Res. Inst. Univ. Tokyo* **76**, 455–478 (2002).
12. Ulmer, P. & Trommsdorff, V. Serpentine stability to mantle depths and subduction related magmatism. *Science* **268**, 858–861 (1995).
13. Rispoli, R. Stress fields about strike-slip faults inferred from stylolites and tension gashes. *Tectonophysics* **75**, T29–T36 (1981).
14. Fletcher, R. C. & Pollard, D. D. Anticrack model for pressure solution surfaces. *Geology* **9**, 419–424 (1981).
15. Aydin, A. & Nur, A. Evolution of pull-apart basins and their scale independence. *Tectonics* **1**, 91–105 (1982).
16. Burnley, P. C. *The Effect of Nonhydrostatic Stress on the Olivine-spinel Transformation in Mg₂GeO₄*. Thesis, Univ. California, Davis (1990).
17. Bureau, H. & Keppler, H. Complete miscibility between silicate melts and hydrous fluids in the upper mantle: experimental evidence and geochemical implications. *Earth Planet. Sci. Lett.* **165**, 187–196 (1999).
18. Agee, C. B. & Walker, D. Static compression and olivine flotation in ultrabasic silicate liquid. *J. Geophys. Res.* **93**, 3437–3449 (1988).
19. Silver, P. G. *et al.* Rupture characteristics of the deep Bolivian earthquake of 1994 and the mechanism of deep-focus earthquakes. *Science* **268**, 69–73 (1995).
20. Hacker, B. R., Peacock, S. M., Abers, G. A. & Holloway, S. D. Subduction factory – 2. Are intermediate-depth earthquakes in subducting slabs linked to metamorphic dehydration reaction? *J. Geophys. Res.* **108**, doi:10.1029/2001JB001129 (2003).
21. Jiao, W., Silver, P. G., Fei, Y. & Prewitt, C. T. Do intermediate- and deep focus earthquakes occur on preexisting weak zones? An examination of the Tonga subduction zone. *J. Geophys. Res.* **105**, 28125–28138 (2000).
22. Christova, C. & Scholz, C. H. Stresses in the Vanuatu subducting slab: A test of two hypotheses. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL017701 (2003).
23. Zhao, D., Mishra, O. P. & Sanda, R. Influence of fluids and magma on earthquakes: seismological evidence. *Phys. Earth. Planet. Inter.* **132**, 249–267 (2002).
24. Zhang, J., Green, H. W. II, Bozhilov, K. N. & Jin, Z.-M. Faulting induced by precipitation of water at grain boundaries in hot subducting oceanic crust. *Nature* (in the press).
25. Stalder, R. & Ulmer, P. Phase relations of a serpentine composition between 5 and 14 GPa: significance of clinohumite and phase E as water carriers into the transition zone. *Contrib. Mineral. Petrol.* **140**, 670–679 (2001).
26. Niu, F., Silver, P. G., Nadeau, R. M. & McEvilly, T. V. Migration of seismic scatterers associated with the 1993 Parkfield aseismic transient event. *Nature* **426**, 544–548 (2003).
27. Bromiley, G. D. & Pawley, A. R. The stability of antigorite in the systems MgO–SiO₂–H₂O (MSH) and MgO–Al₂O₃–SiO₂–H₂O (MASH): The effects of Al³⁺ substitution on high-pressure stability. *Am. Mineral.* **88**, 99–108 (2003).
28. Green, H. W. II & Borch, R. S. A new molten salt cell for precision stress measurement at high pressure. *Eur. J. Mineral.* **1**, 213–219 (1989).
29. Green, H. W., Young, T. E., Walker, D. & Scholz, C. H. Anticrack-associated faulting at very high-pressure in natural olivine. *Nature* **348**, 720–722 (1990).

Acknowledgements We thank V. Trommsdorff for sample material corresponding to that used in ref. 12; J. Zhang for discussions; K. Bozhilov for assistance with microscopy and photography; and F. Forgit for specimen assembly preparation and laboratory assistance. This work was supported by the US National Science Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.W.G. (hgreen@mail.ucr.edu).

The decline and fate of an iron-induced subarctic phytoplankton bloom

Philip W. Boyd¹, Cliff S. Law², C.S. Wong³, Yukihiro Nojiri⁴, Atsushi Tsuda⁵, Maurice Levasseur⁶, Shigenobu Takeda⁷, Richard Rivkin⁸, Paul J. Harrison^{9,10}, Robert Strzepek¹¹, Jim Gower³, R. Mike McKay¹², Edward Abraham², Mike Arychuk³, Janet Barwell-Clarke³, William Crawford³, David Crawford³, Michelle Hale⁸, Koh Harada¹³, Keith Johnson³, Hiroshi Kiyosawa¹⁴, Isao Kudo¹⁵, Adrian Marchetti¹⁶, William Miller¹⁷, Joe Needoba⁹, Jun Nishioka¹⁸, Hiroshi Ogawa¹⁹, John Page³, Marie Robert³, Hiroaki Saito²⁰, Akash Sastri²¹, Nelson Sherry⁹, Tim Soutar³, Nes Sutherland³, Yosuke Taira⁵, Frank Whitney³, Shau-King Emmy Wong³ & Takeshi Yoshimura¹⁸

¹NIWA Centre for Chemical and Physical Oceanography, Department of Chemistry, University of Otago, PO Box 56, Dunedin, 9003, New Zealand

²National Institute of Water and Atmospheric Research, 301 Evans Bay Parade, Greta Point, PO Box 14-901, Kilbirnie, Wellington, New Zealand

³Fisheries and Oceans Canada, Institute of Ocean Sciences, PO Box 6000, Sidney, British Columbia, V8L 4B2 Canada

⁴National Institute for Environmental Studies, Tsukuba, Ibaraki 305-8506, Japan

⁵Ocean Research Institute, University of Tokyo, Nakano, Tokyo 164-8639, Japan

⁶Department of Biology, University of Laval, Quebec, G1K 7P4, Canada

⁷Department of Aquatic Bioscience, University of Tokyo, Bunkyo, Tokyo 113-8657, Japan

⁸Ocean Sciences Centre, Memorial University of Newfoundland, St John's, Newfoundland, A1C 5S7, Canada

⁹Earth and Ocean Sciences, University of British Columbia, 6270 University Blvd, Vancouver, British Columbia, V6T 1Z4, Canada

¹⁰Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

¹¹Department of Chemistry, University of Otago, PO Box 56, Dunedin, 9003, New Zealand

¹²Department of Biological Sciences, Bowling Green State University, Bowling Green, Ohio 43403, USA

¹³National Institute of Advanced Industrial Science and Technology, 1-3-1, Kasumigaseki, Chiyoda-ku, Tokyo 100-8921, Japan

¹⁴Marine Biological Research Institute of Japan, Shinagawa, Tokyo 142-0042, Japan

¹⁵Graduate School of Fisheries Science, Hokkaido University, Hakodate, Hokkaido 041-8611, Japan

¹⁶Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, British Columbia, V6T 1Z4, Canada

¹⁷Dalhousie University, Department of Oceanography, Halifax, Nova Scotia, B3H 4J1 Canada

¹⁸Central Research Institute of Electric Power Industry, Abiko, Chiba 270-1194, Japan

¹⁹Marine Biogeochemistry Laboratory, Ocean Research Institute, The University of Tokyo 1-15-1, Minamidai, Nakano, Tokyo 164-8639, Japan

²⁰Tohoku National Fisheries Research Institute, Shiogama, Miyagi 985-0001, Japan

²¹School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, V8W 3N5, Canada

Iron supply has a key role in stimulating phytoplankton blooms in high-nitrate low-chlorophyll oceanic waters^{1–5}. However, the fate of the carbon fixed by these blooms, and how efficiently it is exported into the ocean’s interior, remains largely unknown^{1–5}. Here we report on the decline and fate of an iron-stimulated diatom bloom in the Gulf of Alaska. The bloom terminated on day 18, following the depletion of iron and then silicic acid, after which mixed-layer particulate organic carbon (POC) concentrations declined over six days. Increased particulate silica export via sinking diatoms was recorded in sediment traps at depths between 50 and 125 m from day 21, yet increased POC export was not evident until day 24. Only a small proportion of the

mixed-layer POC was intercepted by the traps, with more than half of the mixed-layer POC deficit attributable to bacterial remineralization and mesozooplankton grazing. The depletion of silicic acid and the inefficient transfer of iron-increased POC below the permanent thermocline have major implications both for the biogeochemical interpretation of times of greater iron supply in the geological past^{6,7}, and also for proposed geo-engineering schemes to increase oceanic carbon sequestration^{3,8}.

The magnitude of iron supply to the oceans has changed over geological timescales^{6,9}, and the increased export of carbon from iron-stimulated phytoplankton blooms is one of several mechanisms proposed for past reductions in atmospheric CO₂ concentrations⁶. Mesoscale *in situ* iron enrichments have resulted in diatom blooms, demonstrating that phytoplankton growth in high-nitrate low-chlorophyll (HNLC) waters is controlled by the iron supply^{1–5}. The deliberate iron enrichment of HNLC regions has therefore been considered as a potential strategy to mitigate the current rise in atmospheric CO₂ (refs 3 and 8). However, mesoscale iron enrichments have not studied the decline of a bloom^{1–5}, or quantified the associated POC export^{1–5}.

During the subarctic ecosystem response to iron enrichment study (SERIES), we made the first comprehensive time-series measurements of the decline and fate of an iron-induced diatom bloom. During the bloom decline, the mixed-layer POC inventory decreased by 258 mmol m⁻², yet only 18% of this POC was exported to a depth of 50 m, with 8% being detected below the permanent thermocline (~120 m, ref. 10). We can account for at least 69% of this decrease in iron-elevated mixed-layer POC, with the majority due to bacterial remineralization both within and below the mixed layer.

An initial survey on 7 and 8 July 2002 located a site northeast of Ocean Station Papa (OSP, 50° N, 145° W), in the Gulf of Alaska, with HNLC characteristics typical of this region^{10–12}. On 9 July 2002, a 77 km² patch of HNLC waters was enriched with sulphur hexafluoride (SF₆) (>400 fmol l⁻¹) and dissolved iron (>1 nmol l⁻¹,

Fig. 1a). SERIES commenced on 10 July 2002 (day 0), and concluded on 4 August 2002 (day 25). On day 6, a second iron infusion (without SF₆ tracer) raised the dissolved iron to 0.6 nmol l⁻¹ (Fig. 1a). The patch of iron-enriched waters stretched into an ellipsoid along a north–south axis, increased in area to >200 km² by day 13, and to a maximum of ~1,000 km² by days 17/18, at which time a 70–80 μatm drawdown of CO₂ was evident (Supplementary Fig. 1a).

The development of an iron-stimulated bloom was evident from increases in chlorophyll to ~5 mg m⁻³ (Fig. 1b), with a corresponding increase in POC and particulate silica (PSi) concentrations (Fig. 1c). Diatom biomass was initially low, increased exponentially (Fig. 1b), and then decreased around day 18 corresponding to the decline in chlorophyll concentrations (Fig. 1b). The onset of algal iron limitation occurred around day 12, after which photosynthetic competence (F_v/F_m) progressively decreased to the levels observed in the surrounding waters (Supplementary Fig. 1b). The dissolved iron decreased to HNLC concentrations¹¹ by day 12 (Fig. 1a), followed by silicic acid depletion to <2 mmol m⁻³ by day 16 (Fig. 1d), despite the re-supply of silicic acid via lateral entrainment of surrounding waters (see Methods). Microzooplankton grazing mortality on phytoplankton of <202 μm was initially 0.42 d⁻¹, then decreased to 0.09 d⁻¹ by day 6 and remained constant during the bloom evolution (compare with algal growth rates of up to 0.45 d⁻¹ on day 13, data not shown). Mesozooplankton herbivory was highest during the bloom decline (10.8 mmol of carbon per m² per day, compared with 6.0 mmol C m⁻² d⁻¹ in the surrounding waters) with daily rates equivalent to 2% of the POC inventory on day 18 (Fig. 1c). Thus, bloom termination was driven by resource limitation, rather than top-down control¹³.

The decline of the bloom was signalled by rapid decreases in POC and PSi concentrations at the patch centre after day 18 (Fig. 1c) with 75–79% deficits in surface mixed layer inventories by day 25 (Table 1). The bloom decline was also evident from space, with sea-viewing wide field of view sensor (SeaWiFS) images providing

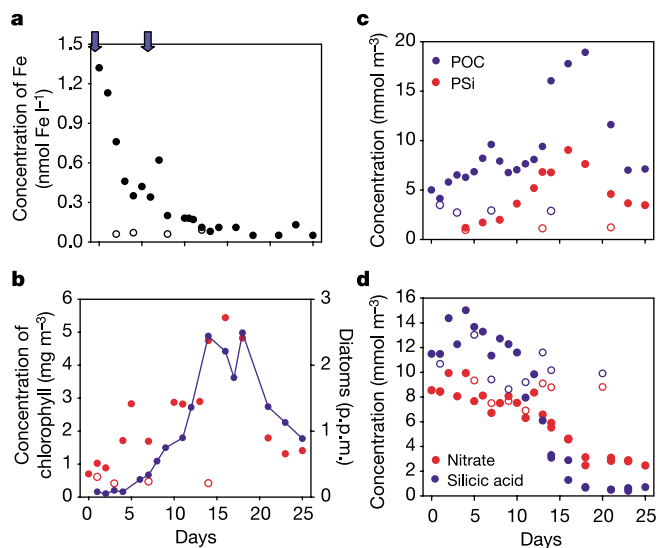


Figure 1 Time series of 'In' (patch centre, solid symbols) and 'Out' (surrounding waters, open symbols) mean mixed layer properties. **a**, Dissolved iron, arrows denote iron infusions. **b**, Chlorophyll (red) and diatom biovolume¹⁴ (blue, 5 m depth). **c**, POC (blue) and PSi (red). **d**, Nitrate (red, 5 m) and silicic acid (blue, 5 m). The standard error ($n = 3$) was: <0.02 nmol l⁻¹ for dissolved iron; smaller than the symbols for nitrate and silicic acid; <0.1 below 1 mg chl m⁻³, and <0.3 for chlorophyll between 1 and 5 mg m⁻³. POC data were derived from 1 m resolved transmissivity profiles and a POC algorithm²⁷ (see Methods).

Table 1 Upper ocean budgets of biogenic particles at the patch centre

Parameter	Depth (m)	POC (mmol m ⁻²)	PSi (mmol m ⁻²)
Accumulation until day 18	0–25	328	200
Inventory on day 25	0–25	70	51
Mixed layer deficit of POC days 18–23	0–25	-258	-149
Fate of mixed-layer POC deficit			
Export flux	at 50	-47 (18%)	-51 (34%)
Undertrapping*	at 50	-11	
Mesozooplankton herbivory†	0–30	-29	
Remineralization to DIC‡	10–30	-90	
Differential remineralization of PSi and POC from trap§	0–50	-41	
PSi dissolution	10–30	-128	
Remineralization based on NH ₄ accumulation¶	0–25	-100	

Water column inventories and fluxes are defined as ('In' minus 'Out'). Loss terms are assigned a negative sign. Between 69% and 99% of the observed mixed layer POC deficit was accounted for. Export efficiency is reported as a percentage, and defined as: (POC export/POC deficit) × 100. For this export calculation, inventories were not corrected for dilution because the patch size was greater (200–1,000 km²) than the source square (for mixed-layer particles intercepted at depth, 50–200 km²) throughout. DOC concentrations changed little between days 18 and 23 and ranged from 1.27 ± 0.03 to 1.34 ± 0.03 mol m⁻² in the patch (0–20 m integrals). Microzooplankton herbivory was not included as a POC loss term because there are no data after day 17. We did not consider data on carbon losses before day 18 because during this period there were no significant differences in the magnitude of loss terms between 'In' and 'Out' waters.

*Potential particle flux (24%) not intercepted by the traps based on a radiochemical trap calibration (see Methods).

†Estimate based on the mean of five grazing experiments on days 18, 19, 20, 22 and 23. The standard error of the mean was 0.6 mmol m⁻² d⁻¹.

‡Derived from increases in the DIC inventory between days 18 and 23 only (Fig. 2c).

§Derived from bacterially mediated differential remineralization of PSi relative to POC¹⁶ at 50 m, and equivalent to a POC loss of 41 mmol m⁻² ((34% minus 18%) × 258 mmol C m⁻²).

||Calculated from excess silicic acid observed between days 18 and 23. Variability in silicic acid concentrations below 30 m precludes an estimate of DIC changes at depth. Remineralization estimates based on PSi and silicic acid of 169 mmol C m⁻² (128 + 41) are higher than that derived from DIC increases (90 mmol C m⁻²), and give a POC loss of 99% (128 + 41 + 29 + 11 + 47).

¶Derived from a 15.1 mmol N m⁻² accumulation of ammonium between days 18 and 23, converted to units of carbon using a Redfield C:N ratio of 6.6.

estimates of bloom areal extent, and homogeneity of algal biomass across the patch (Supplementary Fig. 2). During this decline phase, there were changes in downward particle fluxes (see Methods) below the patch ('In') at 50 to 125 m depth (Fig. 2a, b). Particle fluxes in the surrounding waters ('Out') were relatively constant at 50 m (see Methods), and the 'Out' and 'In' particle fluxes were similar at 50 m from days 1 to 18 (Fig. 2a, b). However, between days 18 and 21, particle abundances intercepted by 50 m 'In' traps increased eightfold, and were dominated both numerically (7:1) and volumetrically (3:1)^{14,15} by diatom cells and aggregates compared with mesozooplankton faecal pellets. From day 21 onwards, 'In' particle fluxes exceeded those for 'Out' traps, with a threefold increase in P*Si* flux at 50 m depth (Fig. 2a). This trend was also observed for the deeper traps. In contrast to P*Si* fluxes, POC fluxes at 50 m rose marginally between days 21 and 23 (Fig. 2b), but then increased to 40 mmol m⁻² d⁻¹ on day 24.5, and were elevated at greater depth (Fig. 2b).

The mixed-layer POC inventory declined by 79% (Table 1), as a result of vertical export (Fig. 2a, b), bacterial solubilization^{16,17}, and ingestion by grazers¹⁸ both within and below the mixed layer. Decreases in mixed-layer POC were not due to physical factors such as vertical advection (Supplementary Fig. 3), or dilution arising from lateral spreading (Supplementary Fig. 2). Only 18% of the decrease in mixed-layer POC was subsequently intercepted by the 50 m trap (Table 1), although the trap may not have captured all of the settling particles¹⁹. Radiochemical calibration of a trap array identical to that deployed in SERIES (see Methods) indicates that these traps potentially undersampled export flux by up to 24% (Table 1), signifying that a maximum 22% of the mixed-layer POC deficit could reach 50 m. Mesozooplankton herbivory of 4.8 mmol

C m⁻² d⁻¹ (see Methods) over six days accounted for 10% of the POC decline (Table 1). Between days 21 and 25, bacterial production rates of 0.3–0.4 mmol C m⁻³ d⁻¹ (upper 6 m only, see Methods) in the patch were double that measured on day 14, and three- to fourfold greater than in the surrounding waters. We used three indirect approaches to estimate bacterial remineralization of particles: changes in dissolved inorganic carbon (DIC) concentrations, selective preservation of opal, and ammonium accumulation (see Methods).

During the decline of the bloom between days 18 and 23, bacterial remineralization of particles¹⁷ was responsible for increased DIC concentrations at depths of 10–30 m (that is, within and below the mixed layer, Fig. 2c), corresponding to concurrent decreases in POC within this depth horizon (Fig. 2d). This increase in DIC was equivalent to ~35% of the mixed-layer POC deficit (Table 1). Bacterial remineralization of POC was also inferred from selective preservation of P*Si* relative to POC^{16,20} in the trap samples from 50 m depth (Table 1). Laboratory radiotracer experiments examining remineralization of diatom detritus have shown that POC is used considerably faster than P*Si* over a timescale of 4–6 days¹⁶. We used the P*Si* intercepted by the 50 m trap to calculate the additional amount of POC remineralized between 0 and 50 m (Table 1). Note that this is a lower bound, because some dissolution of P*Si* will have occurred in the overlying waters, as was confirmed by an increase in silicic acid of 51 mmol m⁻² between days 18 and 23 over the 10–30 m horizon coincident with the DIC increase (Fig. 2c). POC remineralization is a precursor of P*Si* dissolution²⁰, so this increase in silicic acid is equivalent to a DIC increase of 128 mmol m⁻² (using a mixed layer P*Si*:POC molar ratio of 0.4, data not shown). The ammonium accumulation in the upper 25 m of the patch

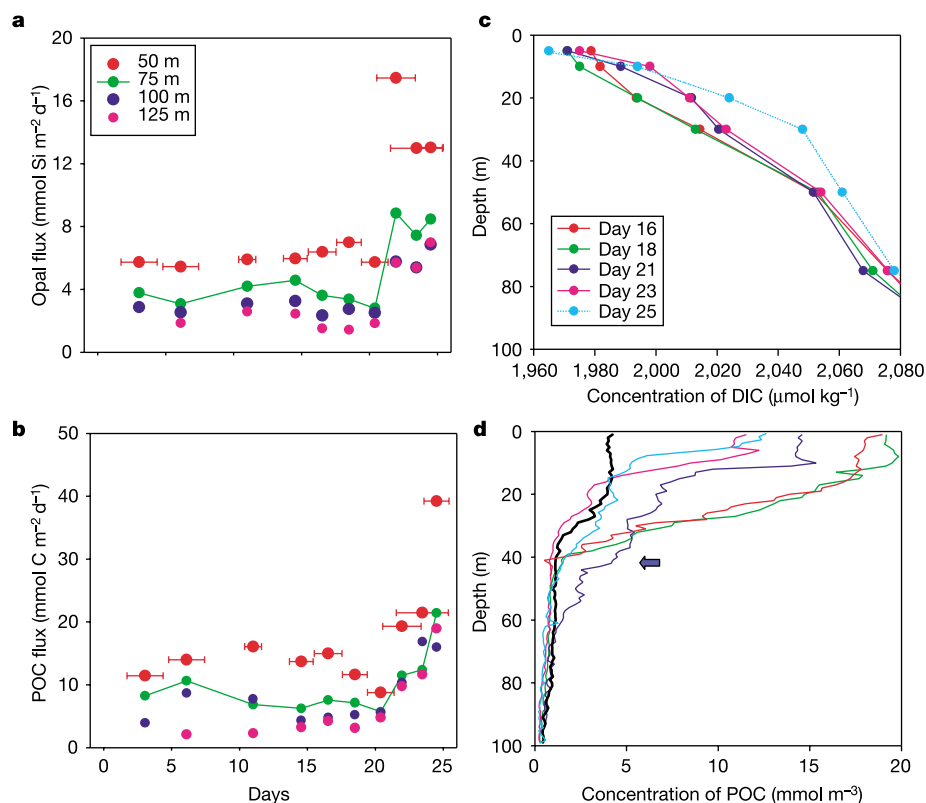


Figure 2 Export and remineralization at the patch centre during the bloom decline. **a**, Downward opal flux intercepted by trap arrays. Symbols denote the deployment midpoint, and horizontal bars the deployment timescales. **b**, POC fluxes from the corresponding deployments as in **a**. **c**, DIC profiles. Standard error bars were always

smaller than symbols. Changes in DIC were not assessed after day 23 owing to the intrusion of denser waters (20–50 m) (Supplementary Fig. 3). **d**, POC profiles for the patch centre²⁷. Colour-coding as in **c**, the black profile represents an 'Out' station (day 14). The arrow denotes sinking and/or stationary particles between 35 and 50 m on day 21.

Table 2 **Calculated Fe:C molar ratios during the SERIES bloom**

	Fe:C	Fe:C ratio	Data source
Added Fe:carbon fixation*	41.3 μmol :1.61 $\pm$ 0.12 mol	2.6×10^{-5} ($\pm 1.9 \times 10^{-6}$) mol mol ⁻¹	Patch centre
Added Fe:POC accumulation†	2.43 tonnes FeSO ₄ ·7H ₂ O: 1,776 $\pm$ 442 tonnes POC	$7.9 \times 10^{-5} \pm 2.0 \times 10^{-5}$	Entire patch†
Added Fe:POC export (50 m)‡	41.3 μmol :0.05 mol	8.3×10^{-4} mol mol ⁻¹	Export signal is integrated over 50–200 km ² ‡
Added Fe:POC export (125 m)‡	41.3 μmol :0.02 mol	2.1×10^{-3} mol mol ⁻¹	Export signal is integrated over 50–200 km ² ‡

The carbon fixation estimate takes into account dilution of the patch, that is, it includes all algal carbon fixation supported by iron enrichment (see Methods). Total iron added during two infusions was 41.3 μmol Fe m⁻² (patch-centre concentrations were raised by 2.25 μmol Fe m⁻³ over 0–15 m (measured 16 h after the iron enrichment began), and by 0.3 μmol Fe m⁻³ over 0–25 m (day 6).

*Based on iron-mediated increases in net primary production of 1.61 ± 0.12 mol C m⁻². Increases in net primary production were greater than decreases in DIC estimated from a carbon budget of the patch centre (1.1 ± 0.07 mol C m⁻², see Methods), due in part to remineralization of POC to DIC (see Fig. 2c).

†Denotes an estimate of POC accumulation based on mean $\pm$ standard deviation of chlorophyll from the SeaWiFS Ocean Colour image on day 19 (Supplementary Fig. 3) multiplied by the carbon:chlorophyll ratio for day 18 at the patch centre (80 g:g) and a mixed layer depth of 25 m.

‡Denotes the range of calculated source-squares (using current speed and algal sinking rate data) for mixed-layer particles that are intercepted at depth by sediment traps³⁰.

between days 18 and 23 was equivalent to a DIC increase of 100 mmol m⁻², in good agreement with the other two indirect methods. Taken together, loss processes and a sampling artefact account for 69% (based on DIC) to 99% (based on P_{Si}) of the fivefold decrease in mixed-layer POC, with bacterial remineralization accounting for the greatest proportion of the decrease (Table 1).

The bloom was initiated within a SF₆-labelled control volume, enabling the construction of a carbon budget for the patch centre (corrected for dilution, see Methods) to provide estimates of Fe:C stoichiometry^{3,8}. The ratio of Fe:C fixed for the patch centre (2.6×10^{-5} , Table 2), is higher than those (2×10^{-6} – 7×10^{-6}) derived from laboratory cultures²¹. Although this ratio was computed for the patch centre, we extrapolated to a Fe:POC accumulation ratio of 7.9×10^{-5} for the entire patch. Budgets and trap data reveal that the vertical transfer efficiency of POC was 18% between the mixed layer and 50 m (Table 1), increasing to ~50% at depths of 50–125 m (Fig. 2b). As with surface waters (Table 1), it is likely that bacterial remineralization¹⁷ was the main process responsible for the attenuation of the bloom export signal below 50 m. Owing to this inefficient vertical transfer of POC from the mixed layer to 125 m (8%; 21 mmol C m⁻²:258 mmol C m⁻²), the ratios for Fe:C export to depth were substantially higher than those for the mixed layer (Table 2). Although a further 23% of mixed-layer iron-elevated POC remained in the surface waters on day 25 (Table 1), the SERIES 'export efficiency' of 8% in our labelled patch is low compared with naturally occurring open-ocean blooms (15–30%, ref. 22) which were sampled in eulerian mode.

SERIES has provided the first detailed evaluation of the decline and fate of an iron-stimulated bloom. Our findings show that although bloom dynamics are controlled by iron supply, bloom termination is also influenced by silicic acid availability. The observed depletion of silicic acid is in contrast to results from modelling simulations of iron-stimulated diatom blooms in the glacial Southern Ocean⁷. Furthermore, silicic acid depletion indicates that an increased iron supply, as occurred during the glacial periods^{6,9}, may not have been the only prerequisite to sustain blooms of siliceous algae. Our study indicates that a concurrent supply of silicic acid would be required for iron-stimulated diatom blooms to persist; this possibility has been reported from the geological record²³. The added Fe:C exported below the permanent thermocline near OSP of 2.1×10^{-3} is a thousandfold higher than that used by geo-engineers to assess the efficiency and cost-effectiveness of iron enrichment as a strategy for mitigating the effects of climate change^{3,8}. Proposed schemes for oceanic iron fertilization to ameliorate anthropogenic increases in atmospheric CO₂ have advocated a strategy of repeated iron enrichments^{3,8}. Our results indicate that the influence of silicic acid depletion may negate the impact of such repeated iron enrichment on diatom stocks, and moreover that inefficient vertical transfer of carbon may limit the effectiveness of iron fertilization as a mitigation strategy. □

Methods

Iron/SF₆ release and mapping

Concurrent addition of acidified dissolved iron (ferrous sulphate, pH 2)² and SF₆ at 7 m depth followed an expanding square pattern over 4.75×4.74 nautical miles, that was modified for lagrangian drift using an uplink radio-buoy²⁴. The drogued ARGOS-GPS drifter buoy at the patch centre subsequently assisted our daily mapping of the patch²⁴. On day 6, a successful iron re-infusion took place in an expanding rectangle of 7.3 by 3.7 nautical miles. As SF₆ reinfusion was not required, mapping allowed real-time discrimination of the patch centre during reinfusion, with cessation of iron supply when the concentration of SF₆ reached <25 fmol l⁻¹.

Correction for entrainment

Temporal changes in concentrations of properties at the patch centre were driven by bloom dynamics, but were influenced concurrently by lateral entrainment of the surrounding waters, causing dilution or supplementation depending on the concentration gradients. Thus, gradients from high concentrations at the patch centre to low concentrations in 'Out' waters (such as POC) resulted in dilution of POC at the patch centre. The reverse concentration gradient (such as for DIC) maintained lateral supply of DIC to the patch centre from surrounding waters. These additional supply/loss terms were accounted for in a carbon budget, by estimating the strain rate of the SF₆-labelled patch²⁵ for days 1–13, and applying the resulting entrainment factor to the 'In'–'Out' gradient for each property. Shipboard incubations, including primary production, were not subject to entrainment effects, and so were not corrected.

Carbon budget

To obtain estimates of carbon fixed per unit iron added, a budget was constructed on the basis of cumulative changes in net primary production (NPP) and DIC concentrations until day 20. Changes in concentrations (or rates) were obtained from 'In' minus 'Out' values for the period when 'In' values exceeded the corresponding 'Out' values (or vice versa for DIC), resulting in each term being summed over slightly different time periods.

Vertical sampling

Sampling took place at the patch centre as defined by daily SF₆ mapping, and water was obtained for fast-repetition-rate fluorometry (FRRF) (ref. 2) (Supplementary Fig. 2), P_{Si} (ref. 26), DIC (ref. 2) and dissolved iron, using a clean vertical pumping system¹¹. All 'Out' stations (>20 km northeast of the patch) were characterized by background SF₆ concentrations. POC data were obtained from an algorithm validated in waters exhibiting a wide range of primary productivity including HNLC (OSP) and mesotrophic waters²⁷, and initial POC concentrations were comparable to those at OSP²⁷. A limited number of mixed-layer POC measurements were available and agreed closely with those from the algorithm²⁷. POC replicates were not available, but seven POC profiles from the patch centre (day 23) exhibited ~7% variation. No replicates for P_{Si} were available. Mesozooplankton herbivory was estimated by the gut pigment technique²⁸ in conjunction with gut evacuation rates²⁸ and carbon:chlorophyll ratios from corresponding days. Key properties, including macronutrients and $F_{\text{d}}/F_{\text{m}}$, were measured on at least two vessels. Whenever possible identical protocols were used (for example, FRRF). Successful intercalibrations (that is ± 10 –15%) were performed before (for example, dissolved iron) or during SERIES (for example, macronutrients).

Trap deployments and trapping efficiency

Particle export fluxes were derived from surface-tethered free-drifting traps for 'In' and 'Out' waters at depths of 50 m, 75 m, 100 m and 125 m for 1.8–3.8-day intervals, with overlap between some deployments and recoveries. The 50 m 'Out' opal fluxes were 6.3, 4.8 and 5.3 mmol m⁻² d⁻¹ for deployments centred on days 3, 6 and 15, respectively. The 50-m 'Out' POC fluxes were 16.0, 8.9, 11.6 mmol m⁻² d⁻¹ for days 3, 6 and 15, respectively. Calibration of the trapping efficiency using ²³⁴Th (ref. 19), of an identical trap array to that used in SERIES, were undertaken in the northwest Pacific at time series station KNOT²⁶. Two calibrations (May 1999) indicated that trapping efficiency was ~76% at 150 m (by comparing ²³⁴Th flux from the vertical profile of ~1,330 d.p.m. m⁻² d⁻¹ with ~1,030 d.p.m. m⁻² d⁻¹ from traps, where d.p.m. is decays per minute). Trapping efficiencies were $>90\%$ at depths of 80–120 m, and at 50 m were in some cases $>100\%$. So a worst-case scenario assumes 24% undertrapping of particles by our traps. The May 1999 trap calibration is applicable as it was conducted in waters with diatom-dominated particle fluxes²⁶, beneath a 30 m mixed layer, and with upper-ocean

currents of $<20 \text{ cm s}^{-1}$, as observed during SERIES. Trapping efficiency is thought to be determined by the characteristics of sinking particles, trap design and upper-ocean hydrodynamics¹⁹. The location of the 'In' traps relative to the patch was monitored using a logging fluorometer 10 m subsurface on the surface-tethered array. Traps remained near the patch centre; on day 23 the array was at the periphery for 36 h before recovery and therefore underestimated vertical export. 'Out' traps were deployed at least 20 km northeast of the patch. Limited replicates were available from trap cups, and the standard error ($n = 3$) for fluxes was $\pm 1.3 \text{ mmol Si m}^{-2} \text{ d}^{-1}$ and $\pm 2.5 \text{ mmol C m}^{-2} \text{ d}^{-1}$ for the 75 and 100 m 'In' traps on day 14.

Bacterial remineralization of particles

We had no information on bacterial growth efficiency, which was required to calculate bacterial carbon demand²⁰: demand = bacterial production $\times$ (1/growth efficiency), published values display a range from 0.1 to 0.7 (ref. 29). Therefore, we employed indirect approaches (changes in DIC concentrations, selective preservation of opal and ammonium accumulation) to estimate particle remineralization.

Received 16 November 2003; accepted 24 February 2004; doi:10.1038/nature02437.

Published online 17 March 2004.

- Coale, K. H. *et al.* A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. *Nature* **383**, 495–501 (1996).
- Boyd, P. W. *et al.* A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature* **407**, 695–702 (2000).
- Buesseler, K. O. & Boyd, P. W. Will ocean fertilization work? *Science* **300**, 67–68 (2003).
- Gervais, F., Riebesell, U. & Gorbunov, M. Y. Changes in primary productivity and chlorophyll *a* in response to iron fertilization in the Southern Polar Frontal Zone. *Limnol. Oceanogr.* **47**, 1324–1335 (2002).
- Tsuda, A. *et al.* A mesoscale iron enrichment in the western Subarctic Pacific induces a large centric diatom bloom. *Science* **300**, 958–961 (2003).
- Martin, J. H. Glacial-interglacial CO_2 change: The iron hypothesis. *Paleoceanography* **5**, 1–13 (1990).
- Matsumoto, K., Sarmiento, J. L. & Brzezinski, M. A. Silicic acid leakage from the Southern Ocean: A possible explanation for glacial atmospheric pCO_2 . *Glob. Biogeochem. Cycles* **16**, doi:10.1029/2001GB001442 (2002).
- Chisholm, S. W., Falkowski, P. G. & Cullen, J. J. Oceans—Discrediting ocean fertilization. *Science* **294**, 309–310 (2001).
- Petit, J. R. *et al.* Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
- Tabata, S. The general circulation of the Pacific Ocean and a brief account of the oceanographic structure of the North Pacific Ocean. Part I—Circulation and volume transports. *Atmosphere* **13**, 133–168 (1975).
- Nishioka, J., Takeda, S., Wong, C. S. & Johnson, W. K. Size-fractionated iron concentrations in the northeast Pacific Ocean: distribution of soluble and small colloidal iron. *Mar. Chem.* **74**, 157–179 (2001).
- LaRoche, J., Boyd, P. W., McKay, R. M. L. & Geider, R. J. Flavodoxin as an *in situ* marker for iron stress in phytoplankton. *Nature* **382**, 802–805 (1996).
- Cullen, J. J. Hypothesis to explain high-nutrient conditions in the open sea. *Limnol. Oceanogr.* **36**, 1578–1599 (1991).
- Strathmann, R. R. Estimating the organic carbon content of phytoplankton from cell volume or plasma volume. *Limnol. Oceanogr.* **12**, 411–418 (1967).
- Komar, P. D., Morse, A. P., Small, L. F. & Fowler, S. W. An analysis of sinking rates of natural copepod and euphausiid fecal pellets. *Limnol. Oceanogr.* **26**, 172–180 (1981).
- Bidle, K. D., Manganelli, M. & Azam, F. Regulation of oceanic silicon and carbon preservation by temperature control on bacteria. *Science* **298**, 1980–1984 (2002).
- Smith, D. C., Simon, M., Alldredge, A. L. & Azam, F. Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature* **359**, 139–142 (1992).
- Dagg, M. Sinking particles as a possible source of nutrition for the large calanoid copepod *Neocalanus cristatus* in the sub-arctic Pacific Ocean. *Deep-Sea Res.* **140**, 1431–1445 (1993).
- Buesseler, K. O., Michaels, A. F., Siegel, D. A. & Knap, A. H. A 3-dimensional time-dependent approach to calibrating sediment trap fluxes. *Glob. Biogeochem. Cycles* **8**, 179–193 (1994).
- Treguer, P. *et al.* The silica balance in the world ocean: a re-estimate. *Science* **268**, 375–379 (1995).
- Sunda, W. G. & Huntsman, S. A. Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Mar. Chem.* **50**, 189–206 (1995).
- Buesseler, K. O. The decoupling of production and particulate export in the surface ocean. *Glob. Biogeochem. Cycles* **12**, 297–310 (1998).
- Harrison, K. G. Role of increased marine silica input on paleo- pCO_2 levels. *Paleoceanography* **15**, 292–298 (2000).
- Law, C. S., Abraham, E. R., Watson, A. J. & Liddicoat, M. Vertical diffusion and nutrient supply to the surface mixed layer of the Antarctic Circumpolar Current. *J. Geophys. Res.* **108**, doi:10.1029/2002JC001604 (2003).
- Abraham, E. R. *et al.* Importance of stirring in the development of an iron-fertilized phytoplankton bloom. *Nature* **407**, 727–730 (2000).
- Honda, M. C. *et al.* The biological pump in the northwestern North Pacific based on fluxes and major components of particulate matter obtained by sediment-trap experiments (1997–2000). *Deep-Sea Res.* **II** **49**, 5595–5625 (2002).
- Bishop, J. K. B. Transmissiometer measurement of POC. *Deep-Sea Res.* **I** **46**, 353–369 (1999).
- Tsuda, A. & Sugisaki, H. In-situ grazing rate of the copepod population in the western subarctic North Pacific during spring. *Mar. Biol.* **120**, 203–210 (1994).
- Rivkin, R. & Legendre, L. Biogenic carbon cycling in the upper ocean: effects of microbial respiration. *Science* **291**, 2398–2400 (2001).
- Siegel, D. A., Granata, T. C., Michaels, A. F. & Dickey, T. D. Mesoscale eddy diffusion, particle sinking, and the interpretation of sediment trap data. *J. Geophys. Res.* **95**, 5305–5311 (1990).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank the officers and crews and scientists on board the *John P. Tully*, *El Puma* and *Kaiyo Maru*. We are grateful to S. Toews for shoreside logistical support. The manuscript was improved by comments from K. Currie, R. Frew, C. Hurd, P. Boyd, D. Hutchins and T. Trull. This study was supported by NSERC Canada as part of the C-SOLAS programme, CFCAS, DFO PERD (Canada), the New Zealand PGSF (Ocean Ecosystems), and the Global Environmental Research Fund from the Ministry of Environment, the Fisheries Agency, and the Central Research Institute of Electric Power Industry research funding (Japan).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.W.B. (p.boyd@niwa.co.nz).

Ferns diversified in the shadow of angiosperms

Harald Schneider^{1,2}, Eric Schuettpelz¹, Kathleen M. Pryer¹, Raymond Cranfill³, Susana Magallón⁴ & Richard Lupia^{5,6}

¹Department of Biology, Duke University, Durham, North Carolina 27708, USA

²Albrecht-von-Haller-Institut für Pflanzenwissenschaften, Abteilung Systematische Botanik, Georg-August-Universität, Untere Karspüle 2, 37073 Göttingen, Germany

³University Herbarium, University of California, 1001 Valley Life Sciences Building #2465, Berkeley, California 94720, USA

⁴Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Circuito Exterior, Anexo al Jardín Botánico, AP 70-233, México DF 04510, Mexico

⁵Sam Noble Oklahoma Museum of Natural History and

⁶School of Geology and Geophysics, University of Oklahoma, Norman, Oklahoma 73072, USA

The rise of angiosperms during the Cretaceous period is often portrayed as coincident with a dramatic drop in the diversity and abundance of many seed-free vascular plant lineages, including ferns^{1–5}. This has led to the widespread belief that ferns, once a principal component of terrestrial ecosystems⁶, succumbed to the ecological predominance of angiosperms and are mostly evolutionary holdovers from the late Palaeozoic/early Mesozoic era. The first appearance of many modern fern genera in the early Tertiary fossil record implies another evolutionary scenario; that is, that the majority of living ferns resulted from a more recent diversification^{7–10}. But a full understanding of trends in fern diversification and evolution using only palaeobotanical evidence is hindered by the poor taxonomic resolution of the fern fossil record in the Cretaceous¹¹. Here we report divergence time estimates for ferns and angiosperms based on molecular data, with constraints from a reassessment of the fossil record. We show that polypod ferns (>80% of living fern species) diversified in the Cretaceous, after angiosperms, suggesting perhaps an ecological opportunistic response to the diversification of angiosperms, as angiosperms came to dominate terrestrial ecosystems.

The extraordinary diversification of angiosperms throughout the Cretaceous and Tertiary, culminating in an estimated 250,000–300,000 living species³, is well known^{1–4,12} owing to the exceptional fossil record of this lineage. The oldest fossils (Early Cretaceous) that can be unequivocally assigned to specific clades of angiosperms^{3,12,13} correspond to the initially diverging lineages resolved in DNA-based phylogenetic analyses^{14–16}. The well-sampled fossil record of subsequently derived angiosperms is also broadly congruent with phylogenetic analyses^{12,14–16}. Ferns, with more than 10,000 living species, are the second largest group of vascular plants⁷. They attained remarkable levels of diversity and abundance from the Carboniferous to the Jurassic—a richness that is well

documented in the geologic record^{9,10,17}—but these levels were shown to have sharply declined in the Cretaceous^{4,5}. This observation of a decline has been directly attributed to the concurrent rise to dominance by angiosperms¹.

Within ferns, strong support has been demonstrated for the monophyly of polypods, a hyperdiverse derived lineage¹⁸ to which the vast majority of Tertiary fern fossils belong. Whereas the virtual absence of unambiguous polypod fossils in the Cretaceous favours a Tertiary radiation, startling new discoveries from the Early Cretaceous of well-preserved polypod sporangia¹⁹ (Neocomian) and roots¹¹ (Albian) seem to imply that the divergence of polypods had begun by the start of the Cretaceous. We aimed to determine which of these two hypotheses—Tertiary versus Cretaceous radiation of polypods—is best supported, on the basis of an independent analysis of DNA sequence data with integrated fossil time constraints.

We obtained DNA sequence data from two chloroplast genes (*rbcL*, *rps4*) for 45 polypods and other fern taxa to reconstruct phylogenetic relationships using bayesian methods. We used the resulting phylogenetic consensus to re-evaluate critical fern fossils (Supplementary Information) and assign them to extant lineages based on morphological synapomorphies. Using these fossils as age constraints, we estimated divergence times with penalized likelihood²⁰ across the fern consensus phylogeny. To evaluate the effects of phylogenetic uncertainty²¹ due to both topological and branch-length estimation error we also estimated divergence times across 1,000 randomly sampled bayesian trees. To allow for consideration of the fern results relative to angiosperms, we assembled a three-gene (chloroplast *rbcL* and *atpB*, and nuclear small-subunit ribosomal DNA) data set consisting of 95 taxa, reconstructed angiosperm relationships, and estimated their divergence times across the angiosperm consensus phylogeny and 1,000 randomly sampled bayesian trees. For angiosperms, we applied fossil age constraints²² (Table 1) in two ways, differing only in the application of the fossil age of the angiosperm crown group. We were either strict (fixed the age at 132 Myr; ref. 22) or relaxed (applied 132 Myr as a minimum age constraint) in our application of this fossil age. Our phylogenetic results for ferns and angiosperms, plotted against the geological timescale, are shown in Fig. 1. Age estimates, means and standard deviations for well-supported nodes (posterior probability ≥ 0.95) are provided in Table 1.

Our phylogenetic analyses of ferns (Fig. 1a) largely confirm relationships observed in prior studies¹⁸, with osmundaceous ferns (*Osmunda*) resolved as the earliest-diverging leptosporangiate lineage followed, in order of divergence, by filmy ferns (*Hymenophyllum*), gleichenioids (*Gleichenia*), schizaeoids (*Lygodium*), water ferns (*Marsilea* + *Salvinia*), tree ferns (*Plagiogyria* to *Dicksonia*) and polypods. Basal within the polypods is a grade of enigmatic genera (*Lonchitis* and *Saccoloma*) and species-poor families leading to a large clade of derived ferns that is composed of two major, species-rich clades—pteridoids and eupolypods. These two clades include 80% of all living fern species. Eupolypods, which comprise 67% of living ferns, consist of two subclades, eupolypods I and eupolypods II. Our divergence time estimates for ferns are mostly older than those implied by the fossil record but show small standard deviations (Table 1). We find that polypod ferns (node a10), derived ferns (node a15) and eupolypods (node a21) began to diversify during the Middle Jurassic, Late Jurassic, and Albian (Early Cretaceous), respectively (Fig. 1a). The major lineages of derived ferns—pteridoids, eupolypods I and eupolypods II (nodes a16, a22 and a25, respectively)—diversified in the Late Cretaceous.

The results of our phylogenetic analyses of angiosperms (Fig. 1b) mostly agree with previously proposed relationships¹⁴. The minor discrepancies between our phylogeny and other published trees are mostly due to taxonomic sampling differences and do not affect our divergence time estimates for angiosperms. Depending on how we

treated the fossil age of the angiosperm crown group (132 Myr; ref. 22), our divergence time estimates for angiosperms varied. If we were strict in our application of this fossil age (that is, fixed the age of node b10) our age estimates were mostly consistent with the fossil record (Fig. 1b, black tree), with the major lineages of angiosperms, such as euasterids I, euasterids II, eurosids II, and core monocots (nodes b46, b40, b55, and b18, respectively) radiating in the Late Cretaceous. If we were relaxed in our application (that is, applied a minimum age constraint to node b10) our age estimates were more in line with those of previous molecular-based studies²³ (Fig. 1b, grey tree), with the major lineages of angiosperms radiating in the Middle Jurassic to Early Cretaceous.

Lineages-through-time plots based on the chronograms presented in Fig. 1a and b document the accumulation of polypod and angiosperm lineages, primarily in the Mesozoic (Fig. 1c). The initiation of the major diversification of polypods, which we estimate at approximately 100 Myr, took place subsequent to the radiation of angiosperms, regardless of whether we fixed the age of the angiosperm crown group, or merely applied a minimum age constraint to this node (Fig. 1c). To the best of our knowledge, this relationship between polypod and angiosperm diversification—polypods evolving in the ‘shadow of angiosperms’—has never before been demonstrated, although it has been suggested^{1,7–9}.

Our present finding of a Cretaceous increase in polypod diversity, based on divergence time estimates, was not recovered in previous investigations of the diversity and abundance of land plants through the Cretaceous and Early Tertiary^{2,4,5}. Most of these other studies, using macrofossil and palynological data, essentially proposed a replacement of seed-free vascular plants by angiosperms, and did not consider specifically the more restricted group of polypods. Recovering an increase in polypod diversity through the Cretaceous, as we did here, is not possible when relying only on the currently available fossil record because of such inherent biases as: (1) the vast majority of palaeobotanical studies have focused on angiosperms rather than ferns; (2) many polypods are epiphytes or occupy other ecological niches that are less likely to be preserved; and perhaps most importantly (3) unlike their living counterparts, fossilized polypod spores are usually plain and unsculptured due to perine detachment, which erases their species diversity in the fossil record. An exclusively molecular approach, however, also has its limitations²⁴. Integrating fossil time constraints together with molecular data is an increasingly powerful tool for identifying historical trends and events across the tree of life²⁵.

Although the proliferation of angiosperms in the Cretaceous is indeed coupled with the decline of many seed-free vascular plant lineages, their radiation also created new ecospace into which certain lineages could diversify. Such an ecological opportunistic response to the spatially more diverse and complex habitats provided by angiosperms, compared to gymnosperms, has been proposed for epiphytic homosporous lycophytes (Lycopodiaceae) using methods related to those used here²⁶, and may also apply to the Cretaceous radiation of polypod ferns. Many modern-day polypods preferentially grow under the low-light canopy conditions present in angiosperm forests. An unconventional photoreceptor, phytochrome 3 (*PHY3*)^{27,28}, recently identified in the polypod fern *Adiantum*, may provide an ecophysiological explanation for much polypod diversification in these shady habitats. *PHY3* is a chimaeric protein with a red/far-red light receptor phytochrome at its amino-terminal end and a blue-light absorbing phototropin at its carboxy-terminal end. Although the light-induced signalling mechanism of *PHY3* is yet to be solved, its capacity to couple the photosensitivity of phytochrome with phototropin kinase activity would allow both red and blue light to function in phototropism and chloroplast movement, thereby conferring a distinct advantage under low-light canopy conditions²⁷. *PHY3* appears to be restricted to polypod ferns and has not been found in any of the basal fern lineages or seed plants²⁷.

The rise of angiosperms undoubtedly initiated fundamental changes in terrestrial ecosystems and set in motion processes that had important consequences for most extant terrestrial biodiversity. Yet, because of ambiguity regarding the timing of angiosperm forest establishment¹, it is not possible for us to confirm that polypods diversified within the dense forest ecospace created by angiosperms or that the chimaeric photoreceptor *PHY3* was a critical innovation necessary for the radiation of polypods. The diversification that we

observe in polypods may be a response to an abiotic stimulus (for example, a decline in CO₂; climate change, such as global warming; increased tectonic activity or other major geological events). Nevertheless, the results of our study point toward the time period to focus on, to test causal hypotheses concerning the radiation of polypods. Potential links between profound biological events, such as the establishment of angiosperm forests and the diversification of polypods, or between these events and large-scale extrinsic factors,

Table 1 Fossil age constraints and molecular age estimates

Node*	Lineage	Fossil age (Myr)†	Molecular age (Myr)		Node*	Lineage	Fossil age (Myr)†	Molecular age (Myr)		Relaxed molecular age (Myr)	
			Estimate‡	Mean ± s.d.§				Estimate‡	Mean ± s.d.§	Estimate‡	Mean ± s.d.§
a01	Euphyllophytes	380.00	380.00	380.00 ± 0.00	b01	Euphyllophytes	380.00	380.00	380.00 ± 0.00	380.00	380.00 ± 0.00
a02	Seed plants	310.00	310.00	310.00 ± 0.00	b02	Monilophytes	354.00	354.00	354.00 ± 0.00	354.00	354.00 ± 0.00
a03	Monilophytes	354.00	354.00	354.00 ± 0.00	b03	Seed plants		329.38	328.96 ± 2.80	334.18	333.10 ± 3.52
a04	Leptosporangiates	280.00	307.08	307.90 ± 8.65	b04			320.19	320.18 ± 1.90	322.57	322.23 ± 2.43
a05		270.00	281.84	283.47 ± 8.99	b05		310.00	310.00	310.00 ± 0.00	310.00	310.00 ± 0.00
a06			218.04	222.99 ± 9.12	b06	Gnetales		184.82	184.71 ± 8.56	184.91	188.94 ± 11.54
a07	Water ferns	137.00	157.98	161.59 ± 13.60	b07			135.63	136.31 ± 9.00	135.72	140.77 ± 11.90
a08	Tree ferns		188.34	191.94 ± 11.18	b08	Conifers		265.47	266.50 ± 11.80	265.52	263.75 ± 12.99
a09		159.00	159.00	159.42 ± 2.66	b09			210.99	212.36 ± 13.33	211.06	209.66 ± 14.16
a10	Polypods	121.00	175.75	181.87 ± 9.48	b10	angiosperms	132.00	132.00	132.00 ± 0.00	251.77	246.36 ± 14.91¶
a11		99.00	159.28	165.34 ± 9.47	b11			130.37	130.32 ± 0.54	232.12	226.27 ± 15.45¶
a12	Lindsaeoids		97.53	103.47 ± 9.75	b12	Nymphaeaceae	121.00	121.00	121.00 ± 0.00	121.00	121.00 ± 0.00
a13			151.42	155.43 ± 9.65	b13			130.78	130.79 ± 0.34	243.22	237.65 ± 15.60¶
a14	Dennstaedtioids		113.93	118.34 ± 12.59	b14	Austrobaileyales		85.12	85.37 ± 6.84	168.02	158.31 ± 27.13¶
a15	Derived ferns	93.50	144.53	148.56 ± 9.44	b15			129.49	129.50 ± 0.42	234.50	229.11 ± 15.61¶
a16	Pteridoids		97.41	100.75 ± 6.72	b16	Monocots		116.47	116.48 ± 2.36	207.08	204.05 ± 13.12¶
a17		37.00	80.22	83.40 ± 6.18	b17		99.00	107.51	107.49 ± 2.63	189.89	187.70 ± 11.34¶
a18			73.93	76.92 ± 6.29	b18	Core monocots		97.39	97.29 ± 3.15	171.09	169.56 ± 10.28
a19		65.00	80.50	83.34 ± 6.32	b19	Commelinids		86.66	87.56 ± 3.57	152.82	153.16 ± 10.10
a20			57.08	59.17 ± 5.23	b20			81.18	82.02 ± 3.58	142.86	143.37 ± 9.72
a21	Eupolypods		104.69	107.29 ± 8.36	b21	Poales	65.00	70.52	71.25 ± 3.65	123.38	124.15 ± 9.11
a22	Eupolypods I		93.61	96.70 ± 7.96	b22	Laurales	105.50	105.50	105.55 ± 0.42	165.94	159.65 ± 21.62
a23			72.73	75.10 ± 6.97	b23	Piperiales		108.05	109.24 ± 3.43	186.32	182.55 ± 16.18¶
a24			47.67	49.57 ± 4.84	b24	Aristolochiaceae	89.00	90.70	92.30 ± 3.71	157.39	153.31 ± 16.33¶
a25	Eupolypods II		94.52	95.55 ± 8.59	b25			69.01	69.69 ± 4.44	117.29	115.38 ± 11.36
a26			55.50	55.22 ± 5.69	b26	Canellales	121.00	121.00	121.00 ± 0.00	157.06	149.65 ± 17.37
a27			40.44	40.21 ± 4.55	b27	Magnoliales	96.00	101.93	102.62 ± 5.21	158.15	152.81 ± 18.47¶
a28			63.64	70.45 ± 9.38	b28	Chloranthaceae	121.00	121.00	121.00 ± 0.00	124.31	128.77 ± 10.56
a29		65.00	69.70	74.13 ± 7.84	b29			128.22	128.06 ± 0.71	229.35	222.92 ± 15.87¶
a30		37.00	54.88	58.17 ± 6.99	b30	Eudicots	121.00	123.31	123.17 ± 0.92	209.59	203.49 ± 16.14¶
					b31		99.00	120.19	120.62 ± 1.06	196.49	193.41 ± 15.60¶
					b32	Core eudicots		116.25	116.68 ± 1.10	180.31	178.97 ± 13.60¶
					b33			100.41	101.44 ± 4.61	154.14	154.14 ± 13.81¶
					b34	Santalales	49.00	87.41	87.58 ± 4.12	131.31	131.68 ± 10.34
					b35	Caryophyllales	83.50	83.50	84.76 ± 2.23	121.75	123.42 ± 9.59
					b36	Asterids		103.50	104.97 ± 1.88	151.06	154.58 ± 12.03¶
					b37	Cornales	85.80	85.80	85.98 ± 0.79	117.42	116.78 ± 17.12
					b38	Ericales	89.00	89.00	89.15 ± 0.75	120.41	122.94 ± 11.28
					b39			97.22	99.27 ± 2.36	141.26	146.06 ± 11.29¶
					b40	Euasterids II		87.98	90.02 ± 3.38	129.54	133.27 ± 12.33¶
					b41			83.61	85.57 ± 3.60	123.64	126.97 ± 12.38¶
					b42			76.96	78.75 ± 3.94	113.31	116.47 ± 11.32
					b43			75.54	78.01 ± 4.63	115.18	116.66 ± 16.08¶
					b44			69.40	71.52 ± 4.79	106.57	107.45 ± 15.79¶
					b45	Apiales	37.00	50.91	52.40 ± 5.02	80.13	80.19 ± 14.38¶
					b46	Euasterids I		91.90	93.81 ± 2.78	132.75	137.63 ± 10.19¶
					b47			78.26	79.85 ± 2.93	111.45	116.38 ± 7.65
					b48	Solanales	33.70	67.56	68.69 ± 3.61	95.88	99.93 ± 7.21
					b49	Lamiales	33.70	53.38	54.92 ± 3.91	75.78	80.15 ± 7.07
					b50	Gentianales	33.70	58.44	59.84 ± 3.62	82.40	86.81 ± 6.45
					b51			114.68	115.13 ± 1.20	173.54	172.53 ± 12.85¶
					b52	Saxifragales	89.00	89.00	89.01 ± 0.30	112.78	108.73 ± 15.01
					b53	Rosids		110.28	110.80 ± 1.29	156.83	157.13 ± 10.17¶
					b54	Myrtales	85.80	85.80	85.81 ± 0.21	104.36	106.07 ± 8.88
					b55	Eurosids II		97.09	97.96 ± 2.71	133.01	133.71 ± 9.09
					b56	Sapindales	65.00	65.00	65.15 ± 0.69	82.12	80.49 ± 7.82
					b57	Malvales	68.00	71.83	72.92 ± 3.65	97.64	98.53 ± 7.38
					b58	Brassicales		55.27	57.86 ± 5.59	76.87	79.92 ± 10.25
					b59	"Eurosids I"		104.49	104.38 ± 1.34	137.92	137.02 ± 7.47
					b60	Fabales		91.60	90.67 ± 4.09	120.61	118.90 ± 7.96
					b61	Fagales	93.50	93.50	93.50 ± 0.00	93.50	93.53 ± 0.38
					b62	Cucurbitales	54.80	77.01	77.25 ± 4.48	99.74	100.54 ± 7.38
					b63	Rosales	65.00	86.69	86.93 ± 4.47	112.76	113.52 ± 8.06

*Node numbers preceded by a and b correspond to Fig. 1a and b, respectively, with lineage names given where applicable.

†Fossil ages (see Supplementary Information) were applied as minimum constraints except at fixed calibration points a01, b01 and b10 (minimum constraint applied to b10 for relaxed analyses).

‡Molecular age estimates are based on penalized likelihood analyses of bayesian consensus trees (Fig. 1a and b).

§Means and standard deviations are based on penalized likelihood analyses of 1,000 randomly sampled bayesian trees; 1,000 age estimates approximated a normal distribution unless indicated.

||Estimated ages for multiple trees were equal to the minimum age constraint.

¶Estimated ages showed a bimodal distribution.

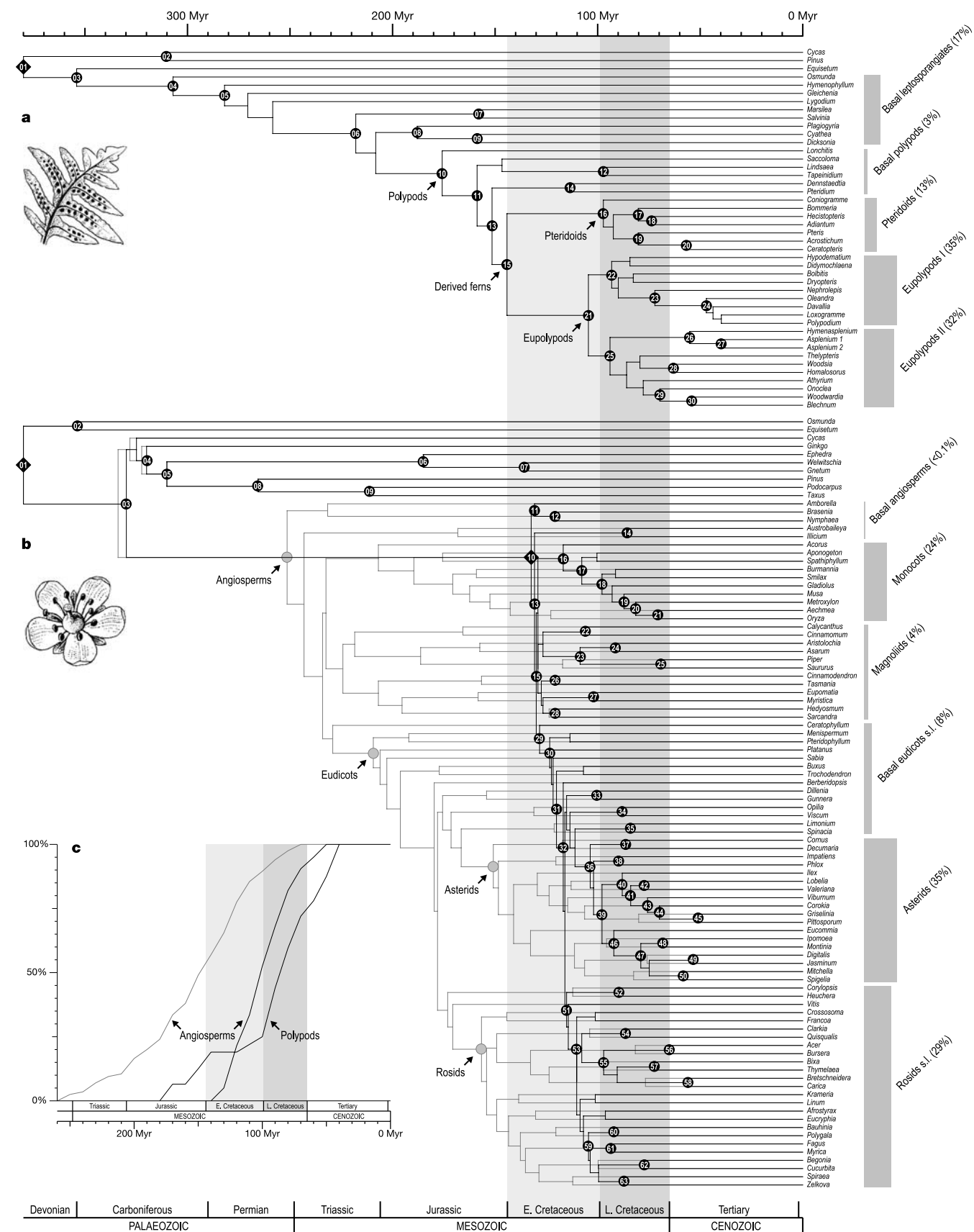


Figure 1 Phylogenetic chronograms of ferns (**a**) and angiosperms (**b**), and proportional lineages-through-time (LTT) plots for angiosperms and polypods (**c**). For angiosperms, the black tree shows strict application of crown group fossil age at b10; grey tree shows relaxed application of this age. Numbered nodes have posterior probabilities ≥ 0.95 ; see Table 1 for lineage names, molecular age estimates and fossil constraints corresponding

to nodes. Nodes in the black angiosperm tree apply also to the grey tree (a few landmarks indicated: b10, b30, b36, b53). Grey boxes on right show per cent species diversity contributed by each lineage. LTT plots for angiosperms (line colour = tree colour) and polypods show number of lineages present (as a proportion of terminals) at sequential time points.

require further exploration to determine their roles in, and possible effects on, the evolution of Cretaceous terrestrial ecosystems. □

Methods

Taxon sampling and DNA sequencing

We assembled two independent data sets (ferns and angiosperms) and used the lycophyte *Huperzia* as the outgroup (not shown in Fig. 1) for both. For ferns, we sampled the chloroplast *rbcl* and *rps4* genes for 45 taxa: 41 leptosporangiate ferns from all major lineages (focusing within the polypods), one horsetail, two seed plants, and the outgroup. For angiosperms, we sampled the chloroplast *rbcl*, *atpB* and nuclear small-subunit ribosomal DNA genes for 95 taxa (mostly a subset from data set in ref. 14): 84 angiosperms, eight gymnosperms, one leptosporangiate fern, one horsetail, and the outgroup. Most DNA sequence data were already available in GenBank; 19 new fern sequences were generated as part of this study following the DNA extraction, amplification and sequencing protocols in ref. 18, and are deposited in GenBank. For fern voucher information and GenBank accession numbers see Supplementary Information.

Phylogenetic analyses

Phylogenetic analyses using a bayesian approach were conducted separately for ferns and angiosperms with MrBayes version 3.0b4 (ref. 29). Each gene in each data set was assigned its own model of evolution (GTR + I + Γ for each gene, determined using a hierarchical likelihood ratio test approach) and analyses were conducted using four chains, run for a total of 10,000,000 generations, with trees sampled from the cold chain every 1,000 generations. The resulting 10,000 trees were plotted against their likelihoods to determine the point where the likelihoods converged on a maximum value, and all 500 trees (500,000 generations) before this convergence were discarded as the 'burn-in' phase. For each data set (ferns and angiosperms), we computed a majority-rule consensus of the remaining 9,500 trees. We used this phylogenetic hypothesis (with average branch lengths), as well as 1,000 trees randomly sampled from among the 9,500 trees, in the subsequent analyses.

Divergence time estimates

Divergence time estimates were obtained through penalized likelihood analyses (truncated Newton algorithm) of the fern and angiosperm consensus phylogenies in r8s version 1.60 (ref. 30). For each data set, fossil age constraints were applied as indicated in Table 1 (for details see Supplementary Information) and the appropriate smoothing value was determined using cross-validation. For angiosperms, fossil age constraints²² (Table 1) were applied in two ways, differing only in the application of the fossil age of the angiosperm crown group (node b10 in Fig. 1b). The first approach was strict (age fixed at 132 Myr; ref. 22) and the second relaxed (minimum age constraint of 132 Myr applied). To evaluate the effects of phylogenetic uncertainty²¹ due to both topological and branch-length estimation error, divergence times were also estimated using penalized likelihood for each of the 1,000 randomly sampled bayesian trees for each data set. Fossil age constraints were applied as described above and individual smoothing values were determined for each tree using cross-validation. The molecular age estimates for each well-supported node (posterior probability ≥ 0.95 ; therefore, 950–1,000 estimates for each node, depending on support) were averaged and the standard deviation calculated.

Lineages-through-time plots

The chronograms resulting from the penalized likelihood analyses of the consensus phylogenies were used to calculate proportional lineages-through-time plots for polypods and angiosperms. Numbers of lineages were tallied at sequential time points (10 Myr intervals) and are presented as proportions (%) of the number of polypod or angiosperm terminals.

Received 5 November 2003; accepted 22 January 2004; doi:10.1038/nature02361.

- Crane, P. R. in *The Origin of Angiosperms and Their Biological Consequences* (eds Friis, E. M., Chaloner, W. G. & Crane, P. R.) 107–144 (Cambridge Univ. Press, Cambridge, 1987).
- Lidgard, S. & Crane, P. R. Angiosperm diversification and Cretaceous floristic trends: A comparison of palynofloras and leaf macrofloras. *Paleobiology* **16**, 77–93 (1990).
- Crane, P. R., Friis, E. M. & Pedersen, K. R. The origin and early diversification of angiosperms. *Nature* **374**, 27–33 (1995).
- Lupia, R., Lidgard, S. & Crane, P. R. Comparing palynological abundance and diversity: Implications for biotic replacement during the Cretaceous angiosperm radiation. *Paleobiology* **25**, 305–340 (1999).
- Nagalingum, N. S., Drinnan, A. N., Lupia, R. & McLoughlin, S. Fern spore diversity and abundance in Australia during the Cretaceous. *Rev. Palaeobot. Palynol.* **119**, 69–92 (2002).
- Niklas, K. J., Tiffney, B. H. & Knoll, A. H. Patterns in vascular land plant diversification. *Nature* **303**, 614–616 (1983).
- Smith, A. R. Comparison of fern and flowering plant distributions with some evolutionary interpretations for ferns. *Biotropica* **4**, 4–9 (1972).
- Lovis, J. D. Evolutionary patterns and processes in ferns. *Adv. Bot. Res.* **4**, 229–415 (1977).
- Rothwell, G. W. in *Pteridology in Perspective* (eds Camus, J. M., Gibby, M. & Johns, R. J.) 395–404 (Royal Botanic Gardens, Kew, 1996).
- Collinson, M. E. in *Pteridology in Perspective* (eds Camus, J. M., Gibby, M. & Johns, R. J.) 349–394 (Royal Botanic Gardens, Kew, 1996).
- Schneider, H. & Kenrick, P. An Early Cretaceous root-climbing epiphyte (Lindsaeaceae) and its significance for calibrating the diversification of polypodiaceous ferns. *Rev. Palaeobot. Palynol.* **115**, 33–41 (2001).
- Dilcher, D. L. Paleobotany: Some aspects of non-flowering and flowering plant evolution. *Taxon* **50**, 697–711 (2001).

- Friis, E. M., Pederson, K. R. & Crane, P. R. Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. *Nature* **410**, 357–360 (2001).
- Soltis, D. E. *et al.* Angiosperm phylogeny inferred from 18S rDNA, *rbcl*, and *atpB* sequences. *J. Linn. Soc. Bot.* **133**, 381–461 (2000).
- Soltis, P. S., Soltis, D. E. & Chase, M. W. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402**, 402–404 (1999).
- Qiu, Y.-L. *et al.* The earliest angiosperms: Evidence from mitochondrial, plastid and nuclear genomes. *Nature* **402**, 404–407 (1999).
- Skog, J. E. Biogeography of Mesozoic leptosporangiate ferns related to extant ferns. *Brittonia* **53**, 236–269 (2001).
- Pryer, K. M. *et al.* Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* **409**, 618–622 (2001).
- Deng, S. Ecology of the Early Cretaceous ferns of Northeast China. *Rev. Palaeobot. Palynol.* **119**, 93–112 (2002).
- Sanderson, M. J. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Mol. Biol. Evol.* **19**, 101–109 (2002).
- Pagel, M. & Lutzoni, F. in *Biological Evolution and Statistical Physics* (eds Lässig, M. & Valleriani, A.) 148–161 (Springer, Berlin, 2002).
- Magallón, S. & Sanderson, M. J. Absolute diversification rates in angiosperm clades. *Evolution* **55**, 1762–1780 (2001).
- Wikström, N., Savolainen, V. & Chase, M. W. Evolution of the angiosperms: Calibrating the family tree. *Proc. R. Soc. Lond. B* **268**, 2211–2220 (2001).
- Benton, M. J. & Ayala, F. J. Dating the Tree of Life. *Science* **300**, 1698–1700 (2003).
- Soltis, P. S., Soltis, D. E., Savolainen, V., Crane, P. R. & Barraclough, T. G. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. *Proc. Natl Acad. Sci. USA* **99**, 4430–4435 (2002).
- Wikström, N. & Kenrick, P. Evolution of Lycopodiaceae (Lycopodiaceae): Estimating divergence times from *rbcl* gene sequences by use of nonparametric rate smoothing. *Mol. Phylogenet. Evol.* **19**, 177–186 (2001).
- Kawai, H. *et al.* Responses of ferns to red light are mediated by an unconventional photoreceptor. *Nature* **421**, 287–290 (2003).
- Smith, H. Phytochromes and light signal perception by plants—an emerging synthesis. *Nature* **407**, 585–591 (2000).
- Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
- Sanderson, M. J. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301–302 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank the Duke Biology systematics discussion group, especially P. Manos, for suggestions; F. Lutzoni, N. Nagalingum and A. R. Smith for comments on the manuscript; and M. Skakuj for the thumbnail sketches included in Fig. 1. This work was supported in part by grants from the National Science Foundation to H.S., K.M.P., R.C. and R.L.; by the Deep Time Research Coordination Network (NSF); and by the A.W. Mellon Foundation Fund to Duke University for Plant Systematics.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to K.M.P. (pryer@duke.edu). Nucleotide sequences newly determined here have been deposited in GenBank under the accession numbers AY459153–AY459171.

Adaptation to natural facial categories

Michael A. Webster, Daniel Kaping, Yoko Mizokami & Paul Duhamel

Department of Psychology, University of Nevada, Reno, Nevada 89557, USA

Face perception is fundamentally important for judging the characteristics of individuals, such as identification of their gender, age, ethnicity or expression. We asked how the perception of these characteristics is influenced by the set of faces that observers are exposed to. Previous studies have shown that the appearance of a face can be biased strongly after viewing an altered image of the face, and have suggested that these after-effects reflect response changes in the neural mechanisms underlying object or face perception^{1–5}. Here we show that these adaptation effects are pronounced for natural variations in faces and for natural categorical judgements about faces. This

suggests that adaptation may routinely influence face perception in normal viewing, and could have an important role in calibrating properties of face perception according to the subset of faces populating an individual's environment.

Physical differences among faces are small compared with the range of variation in many classes of non-face objects, yet observers are highly sensitive to the subtle stimulus differences distinguishing faces. A special capacity for face perception might reflect specialized neural subsystems^{6–8} and expertise developed through extensive experience with faces^{9,10}. Adaptation to faces transfers across differences in the size^{2,3}, retinal position² and orientation^{4,5} of face images, and thus may alter sensitivity directly at these specialized levels. In this study we do not address the site of the sensitivity change, but focus on whether these changes can arise in natural viewing. Even if the after-effects of adaptation reflect more peripheral and general stages of the visual system, the adaptation could be involved centrally in shaping the properties and dynamics of face perception, if the variations among actual faces are large enough to induce different states of adaptation. To test this, we examined adaptation to actual images of faces, and how this affected the types of judgements that observers normally make when viewing faces. These judgements typically involve classifying a face image according to a specific trait or individual identity, and may involve categorical perception (perceiving the stimulus continuum in terms of relatively discrete categories) for such characteristics^{11–13}. Our measurements do not test for categorical perception of faces, but instead test whether the stimulus boundaries that observers choose for different face categories can be shifted by prior adaptation.

Stimuli were formed by morphing between pairs of face images to produce a finely graded series of images spanning the two original faces. The face pairs differed in gender, ethnicity (Japanese or Caucasian) or expression (Fig. 1). Observers made forced-choice responses to categorize images along the continuum (for example, responding as to whether a face from a gender morph appeared 'female' or 'male'). These responses were used to estimate the category boundary at which either response was equally likely. We then asked how these boundaries shifted after observers viewed either of the original face images defining the categories.

The boundary for gender represents an androgynous image intermediate to the female and male exemplars, and could be set consistently by observers (relative to the magnitude of the shifts induced by adaptation). However, after adapting to a male face, the previously ambiguous image appeared distinctly female, and thus the image that now appeared neutral was shifted towards the male image (Fig. 2a). Conversely, adaptation to the female face induced the opposite changes. Differences in the gender boundaries after adaptation were assessed for each individual subject with a two-way analysis of variance (adapted face (male versus female) by test morph (three face pairs)). For all five observers, the mean settings differed significantly for male or female adaptation ($F_{1,18} \geq 23.36$, $P < 0.001$), but did not differ for the specific face pair tested ($F_{2,18} \leq 1.84$, not significant (NS)) and did not show a significant interaction between adapting condition and face pair ($F_{2,18} \leq 1.08$, NS). We found similar after-effects for each of the other face categories we examined. For example, adaptation to a Japanese or Caucasian face again significantly biased the perceived neutral image in the sequence between them ($F_{1,18} \geq 56.1$, $P < 0.001$) (Fig. 2c). These biases were large perceptually (Fig. 3), and are consistent with a shift in the perceived neutral point towards the adapting image, because the adapting face itself appears more neutral. That is, the adaptation may re-centre the 'face space' relative to the adapting facial configuration.

Could observers adapt to the stimulus differences defining traits like gender in the male and female face pairs or only to the differences defining individual identity? To assess this we adapted subjects to a collection of ten male or ten female faces, each

displayed as a random sequence of faces changing every 250 ms during the adaptation intervals. We then measured the gender boundary along male–female image arrays for face images that were not part of the adapting set (Fig. 2b). Once again the category boundaries were significantly biased by adaptation ($F_{1,18} \geq 8.55$; $P < 0.01$). This suggests that the set of faces shared common characteristics plausibly related to their common gender, and that these induced adaptation to the attribute of gender, which transferred to the perception of new individuals. Similarly, adaptation to distortions that alter the rated attractiveness of one set of faces influences the perceived attractiveness of a different set of faces⁴.

Judgements of ethnicity and gender reflect distinctions between individuals. We also examined whether adaptation influenced how variations within a single face are categorized, by measuring the perceived expression of a face. Facial expressions involve an innate and universal pattern of behaviour, and are perceived in terms of a small number of relatively discrete categories (for example, angry, happy or surprised)¹⁴. We morphed between different pairs of expressions in a single face¹⁵, and then determined the stimulus boundary between these pairs before or after adapting to the original expression. The after-effects of adaptation were consistent with those we found for gender and ethnicity (Fig. 2d). For example, in a happy–angry morph, previous adaptation to the angry or happy face induced clear differences in the expression boundary ($F_{1,6} \geq 9.71$, $P < 0.05$). Similar shifts in the neutral point after adaptation occurred in morphs between disgust–surprise ($F_{1,6} \geq 18.54$, $P \leq 0.01$) or fear–contempt ($F_{1,6} \geq 11.78$; $P < 0.05$).

In the course of these measurements we also discovered large individual differences in the category boundaries chosen during the pre-adaptation settings, and these appeared to be related to the categories to which the observers themselves belonged. This is

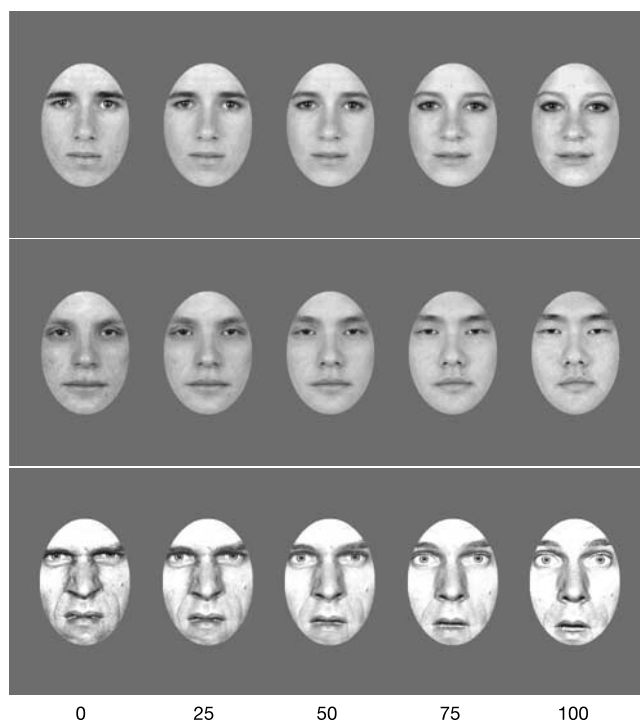


Figure 1 Examples of face pairs for the dimensions of gender (top row), ethnicity (middle row), or expression (bottom row). Morph arrays consisted of 100 images spanning the two original faces, which were assigned a value of 0 or 100. The morph level (for example, 25, 50, or 75) therefore corresponded to how far along the sequence the image fell between the two originals. Category boundaries corresponded to the image level chosen as the neutral point within the morph sequence.

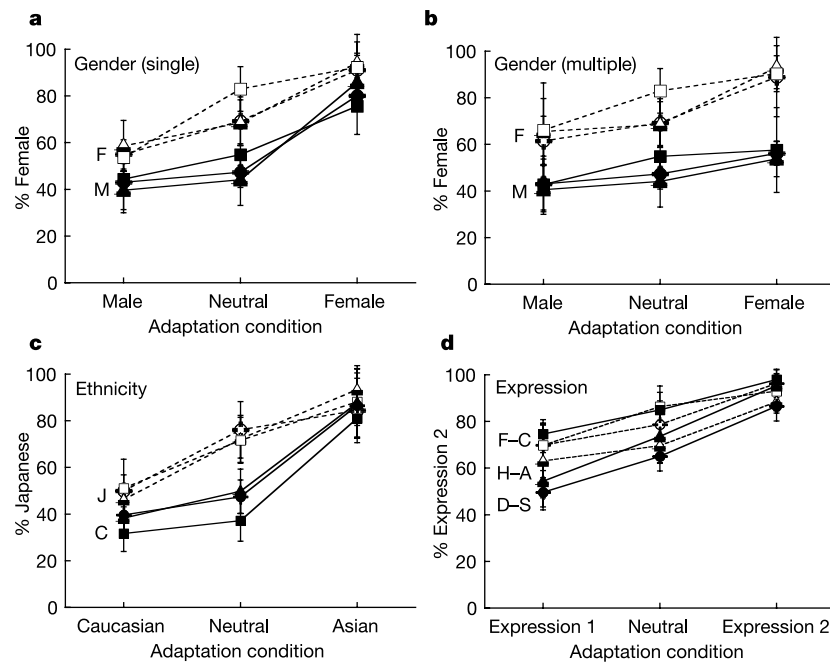


Figure 2 Category boundaries before or after adaptation. Each plot shows mean boundaries (± 1 standard error) set by two observers for three face pairs, before or after viewing images from the two categories defining the morphs. **a**, Settings by a female (F) or male (M) for gender morphs. **b**, Settings by the same observers after adapting to a

series of different male or female faces. **c**, Settings by a Japanese (J) or Caucasian (C) observer for ethnicity morphs. **d**, Boundaries selected by a Caucasian female (open symbols) and male (filled symbols) for morphs between happy–angry (H–A, triangles), disgust–surprise (D–S, circles) or fear–contempt (F–C, squares).

evident for the gender settings for the female and male subject in Fig. 2a. Females and males tended to choose gender boundaries in a male–female continuum that were shifted towards their own gender. We verified this for a set of 28 undergraduate students who set the gender boundaries under neutral adaptation. The average boundary for females (mean = 72, $n = 15$) and males (mean = 39, $n = 13$) (see Fig. 1 legend for an explanation of the mean value) differed by 33 steps in the image arrays, a difference that was again highly significant ($t = 14.08$, degrees of freedom (d.f.) = 26, $P < 0.001$) and perceptually obvious (Fig. 4a). To test whether this was due to a difference in decision strategies for the two genders (for example, females choosing the female side of a gender–neutral region and males choosing the male side of the same region) we repeated the measurements for a new group of ten female and nine male observers who were all instructed to respond to whether the face appeared ‘female’ or ‘not female’ (Fig. 4a). These settings did not differ from the previous group ($t = 0.81$, d.f. = 23, NS for females; $t = 0.81$, d.f. = 20, NS for males) while the differences between females and males within this second group were again large (mean = 68 versus 36; $t = 5.46$, d.f. = 17, $P < 0.001$). These differences suggest that observers may generally be more sensitive to how a face differs from their own category, although the extent to which this reflects differences in perception or criteria remains uncertain.

We next tested whether similar differences in category boundaries might occur for observers that differ in ethnicity. Boundary images for Japanese–Caucasian morphs were obtained for 38 Caucasian undergraduates in Reno, Nevada, and for 42 Japanese exchange students who had recently arrived for study in Reno. The latter group was tested within two weeks of their departure from Japan. The mean boundaries chosen were again very different for the two groups (67 versus 42; $t = 9.80$, d.f. = 78, $P < 0.001$) and were consistent with an increased sensitivity to how the faces differed from the observer’s own ethnicity (Fig. 4b).

The differences between the Japanese and Caucasian students prompted us to test for signs of adaptive adjustments outside the

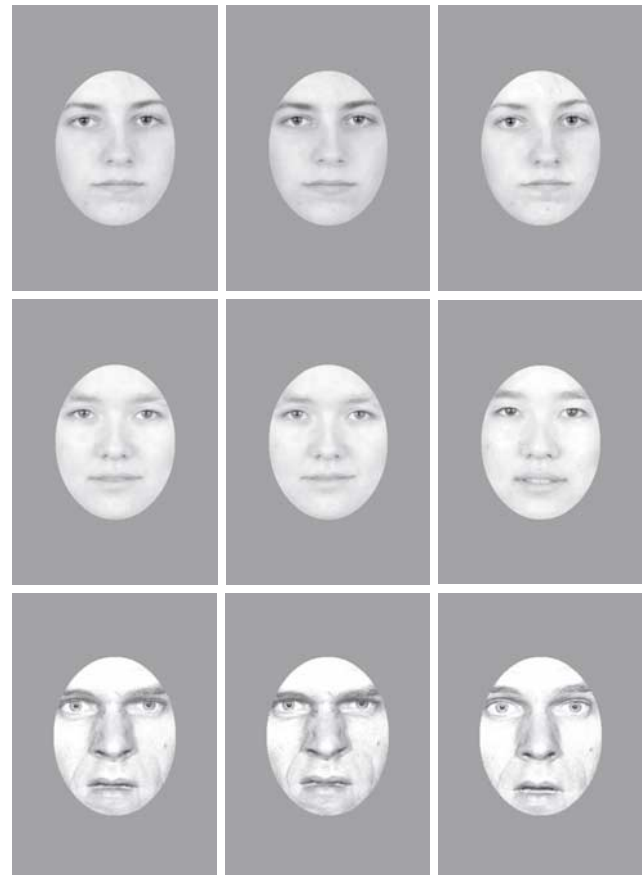


Figure 3 Examples of the face images chosen as category boundaries. Rows show boundary images before (left column) or after adaptation to gender (top row: male adapted, middle; female adapted, right), ethnicity (middle row: Caucasian adapted, middle; Japanese adapted, right), or expression (bottom row: disgust adapted, middle; surprise adapted, right).

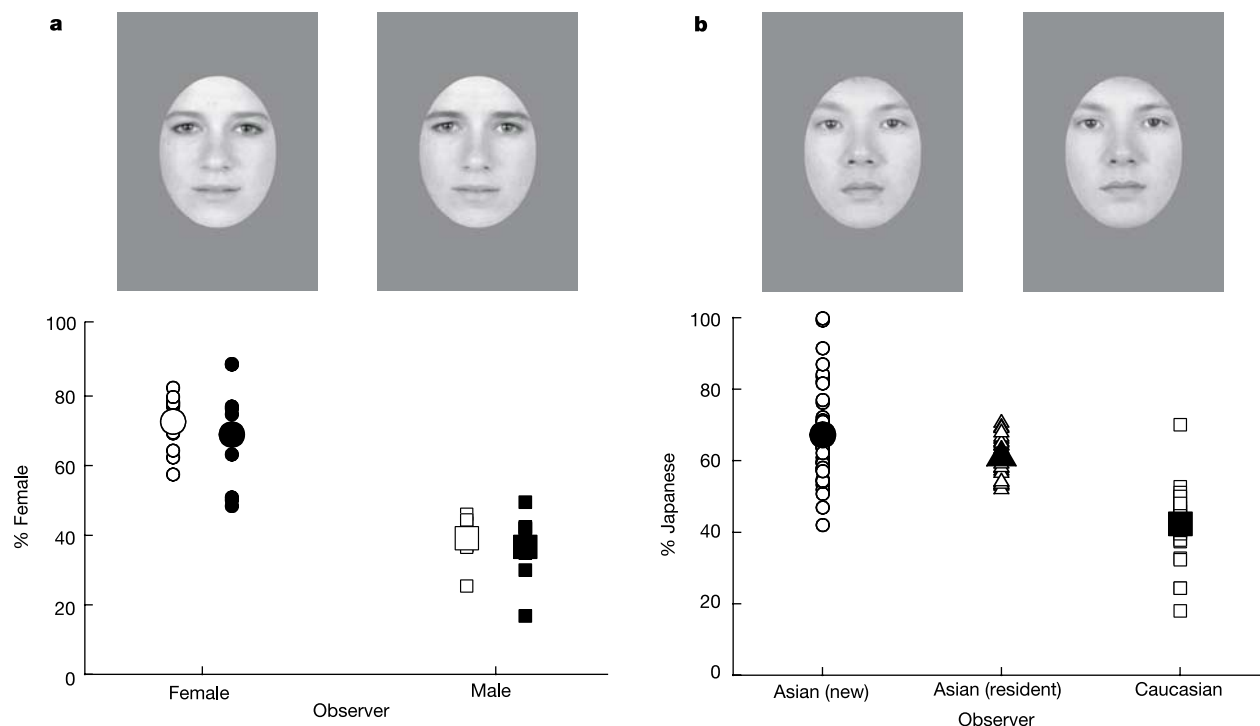


Figure 4 Category boundaries selected by different categories of observers. **a**, Gender boundaries. Open symbols show settings by females (circles) or males (squares) based on a male-female response criterion. Faces corresponding to the group means (enlarged symbols) are shown in the images above. Filled symbols show settings when observers set

the gender boundary based on a female or not-female criterion. **b**, Ethnicity boundaries chosen by Asian students that had newly arrived in Reno, Nevada (circles) or were resident for at least one year (triangles), or for Caucasian students (squares). Faces show the mean images for the newly arrived Asian or for the Caucasian students.

controlled conditions of the laboratory, which could arise if observers are exposed to a change in the sample of faces encountered in their routine daily activities. To test this we measured ethnicity boundaries for a second group of 32 Asian and predominantly Japanese students who had been resident in the US for at least 1 yr. Their settings (mean = 61) were modestly but significantly shifted towards the neutral points for the Caucasian students (Fig. 4b), as confirmed by a *t*-test and (because the samples had very different variances) by a nonparametric Wilcoxon signed ranks test ($Z = -1.81$; $P < 0.05$ one-tailed). Questionnaires administered to a subset from this group revealed that their settings were also negatively correlated with the length of their stay in the United States ($r = -0.57$, $n = 20$, $P < 0.01$), while positively correlated with self-reports of the percentage of time that they spent with individuals of the same ethnicity ($r = 0.59$, $n = 20$, $P < 0.01$). That is, the perceptual boundaries tended to shift more the longer these students were in the United States and the more they interacted with Caucasian individuals. A complete shift may be unlikely, because the gender differences suggest that observers may remain more sensitive to how faces differ from their own facial categories. The role of contact in the 'other race effect' for face perception has been widely investigated but remains poorly understood^{16,17}. Our short-term adaptation effects suggest that one component of long-term contact with a new population may include actual changes in the appearance of faces induced by adaptation.

Variability among faces is obviously a fundamentally important source of information about individual and group identities and the long-term (for example, age) and short-term (for example, emotional) states of an individual. Yet because this variation is not random or uniform we are all exposed to a different diet of faces. Our results suggest that these natural stimulus variations are potentially large enough to induce different states of adaptation in observers, states that may influence strongly how faces are

perceived and interpreted. Adaptation is thought to facilitate efficient coding of low-level stimulus features by normalizing visual responses to the average stimulus levels in scenes, and aids perceptual constancy by discounting variations in both the environment and the observer. Adapting face perception to the average characteristics of the distribution of faces that an individual encounters may similarly be important for calibrating the visual processes encoding faces. □

Methods

Stimuli were frontal-view images of individual faces taken from the JACNeuF and JACFEE image set of ref. 18. The images were converted to greyscale and cropped by a uniform oval (420 by 630 pixels) that preserved the internal features of the face while removing differences in external features of the head, such as hair shape. Pairs of images were then morphed with the program Gryphon Morph, with the morphing sequence stored as 100 frames of a movie file. On each trial an individual frame from the continuum was shown for 500 ms on a computer screen. The face subtended roughly 10° vertical and was displayed on a uniform grey background. Observers used a button box to make a two-alternative forced-choice response to classify the image into one of two categories (corresponding to morphs between two genders, ethnicities, or expressions). Successive images were then varied in two randomly interleaved staircases (that is, with the morph level presented on each trial stepped up or down depending on the response for the preceding trial) of 12 reversals each to estimate the stimulus level for each category boundary. On adaptation trials, subjects first viewed either of the two original images for 180 s, and then for 5 s between each test presentation until the staircases completed. Test and adapting images were separated by 250-ms intervals during which the screen remained uniform grey. In pre-adaptation trials, the uniform background was also shown during the adaptation intervals. Observers were asked to attend to the images throughout but were not otherwise given specific fixation instructions. Four repeated settings were made for three different face pairs for each stimulus category. Five subjects were tested for each condition, and included two of the authors (D.K. and Y.M.) and three naive observers who were unaware of the specific aims of the experiment. Procedures were approved by the Institutional Review Board of the University of Nevada and informed consent was obtained for all subjects.

Received 11 November 2003; accepted 5 February 2004; doi:10.1038/nature02420.

1. Webster, M. A. & MacLin, O. H. Figural after-effects in the perception of faces. *Psychon. Bull. Rev.* **6**, 647–653 (1999).
2. Leopold, D. A., O'Toole, A. J., Vetter, T. & Blanz, V. Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neurosci.* **4**, 89–94 (2001).

3. Zhao, L. & Chubb, C. F. The size-tuning of the face-distortion aftereffect. *Vision Res.* **41**, 2979–2994 (2001).
4. Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G. & Nakayama, K. Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychol. Sci.* **14**, 558–566 (2003).
5. Watson, T. L. & Clifford, C. W. G. Pulling faces: An investigation of the face-distortion aftereffect. *Perception* **32**, 1109–1116 (2003).
6. Moscovitch, M., Winocur, G. & Behrmann, M. What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *J. Cogn. Neurosci.* **9**, 555–604 (1997).
7. Farah, M. J., Wilson, K. D., Drain, M. & Tanaka, J. N. What is “special” about face perception? *Psychol. Rev.* **105**, 482–498 (1998).
8. Kanwisher, N. Domain specificity in face perception. *Nature Neurosci.* **3**, 759–763 (2000).
9. Diamond, R. & Carey, S. Why faces are and are not special: An effect of expertise. *J. Exp. Psychol. Gen.* **115**, 107–117 (1986).
10. Gauthier, I. & Tarr, M. J. Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Res.* **37**, 1673–1682 (1997).
11. Etcoff, N. L. & Magee, J. J. Categorical perception of facial expression. *Cognition* **44**, 227–240 (1992).
12. Beale, J. M. & Keil, F. C. Categorical effects in the perception of faces. *Cognition* **57**, 217–239 (1995).
13. McKone, E., Martini, P. & Nakayama, K. Category perception of face identity in noise isolates configural processing. *J. Exp. Psychol. Hum. Percept. Perform.* **27**, 573–599 (2001).
14. Ekman, P. Strong evidence for universals in facial expression: A reply to Russell’s mistaken critique. *Psychol. Bull.* **115**, 268–287 (1994).
15. Young, A. et al. Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition* **63**, 271–313 (1997).
16. Levin, D. T. Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *J. Exp. Psychol. Gen.* **129**, 559–574 (2000).
17. Furl, N., Phillips, P. J. & O’Toole, A. J. Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis. *Cogn. Sci.* **26**, 797–815 (2002).
18. Matsumoto, D. & Ekman, P. *Japanese and Caucasian Facial Expressions of Emotion (JACFEE) and Neutral Faces (JACNeuF)* (Department of Psychology, San Francisco State Univ., San Francisco, 1988).

Acknowledgements This work was supported by the National Eye Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.A.W. (mwebster@unr.nevada.edu).

A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice

Zhi-yong Yang^{1*}, Wing-pui Kong^{1*}, Yue Huang¹, Anjeanette Roberts², Brian R. Murphy², Kanta Subbarao² & Gary J. Nabel¹

¹Vaccine Research Center, NIAID, National Institutes of Health, Building 40, Room 4502, MSC-3005, 40 Convent Drive, Bethesda, Maryland 20892-3005, USA

²Laboratory of Infectious Diseases, NIAID, National Institutes of Health, MSC-8007, 50 South Drive, Bethesda, Maryland 20892-8007, USA

* These authors contributed equally to this work

Public health measures have successfully identified and contained outbreaks of the severe acute respiratory syndrome (SARS) coronavirus (SARS-CoV)^{1–5}, but concerns remain over the possibility of future recurrences. Finding a vaccine for this virus therefore remains a high priority. Here, we show that a DNA vaccine encoding the spike (S) glycoprotein of the SARS-CoV induces T cell and neutralizing antibody responses, as well as protective immunity, in a mouse model. Alternative forms of S were analysed by DNA immunization. These expression vectors induced robust immune responses mediated by CD4 and CD8 cells, as well as significant antibody titres, measured by enzyme-linked immunosorbent assay. Moreover, antibody responses in mice vaccinated with an expression vector encoding a form of S that includes its transmembrane domain elicited neutralizing antibodies. Viral replication was reduced by more than six orders of magnitude in the lungs of mice vaccinated with these S plasmid

DNA expression vectors, and protection was mediated by a humoral but not a T-cell-dependent immune mechanism. Gene-based vaccination for the SARS-CoV elicits effective immune responses that generate protective immunity in an animal model.

The SARS-CoV emerged in Asia as a highly aggressive pathogen that can be lethal in adults and the elderly^{1–6}. Although its genetic organization is similar to other coronaviruses, recent phylogenetic studies suggest that it may be most closely related to type II coronaviruses^{7–9}. Previously described coronaviruses have not usually induced lethal disease in humans, but their pathogenicity in domestic animal species has been well documented¹⁰, and experimental vaccines developed for animals have provided insight into mechanisms of protective immunity^{11,12}. Studies of the immune response to coronaviruses suggest that both cell-mediated and humoral immunity contribute to long-term protection^{13,14}.

To discover the gene products and mechanisms of protective immunity relevant to SARS-CoV, two sets of cDNAs encoding the SARS-CoV S glycoproteins were prepared using modified codons to optimize expression and to minimize recombination with endogenous coronaviruses. Because coronaviruses assemble in the compartment between the endoplasmic reticulum (ER) and Golgi apparatus¹⁰, and the S leader may direct it to the ER, the native leader sequence was retained in one set of vectors (Fig. 1a) and replaced in another set with a leader sequence derived from the interleukin-2 gene. Expression was not significantly altered by this leader sequence substitution (data not shown), and it was not studied further. Two S carboxy-terminal mutants, one that truncated the cytoplasmic domain (Δ CD) and another that deleted the transmembrane and cytoplasmic regions (Δ TM), were prepared, and expression of these cDNAs by a mammalian expression vector suitable for human vaccination was confirmed (Fig. 1b).

The plasmids encoding these modified S glycoproteins were analysed for their ability to elicit antiviral immunity after intramuscular injection in BALB/c mice. Injection of S, Δ TM and Δ CD expression vectors induced a substantial immune response. A marked increase was observed in the number of SARS-CoV S-specific CD4 T-cell immune responses (Fig. 2a), as measured by intracellular cytokine staining for interferon- γ (IFN- γ) and tumour necrosis factor- α (TNF- α). In addition, substantial SARS-CoV S-specific CD8 cellular immunity was detected at levels at least sevenfold above the background response. Humoral immunity was initially

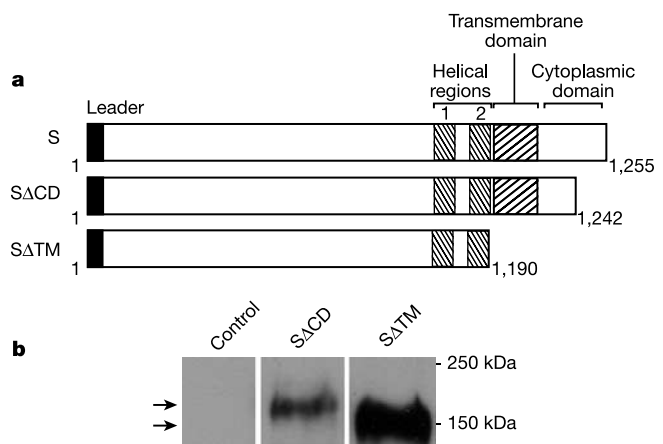


Figure 1 Schematic representation of SARS-CoV glycoprotein cDNAs and expression of recombinant proteins. **a**, The structure of the cDNAs used. **b**, Expression of these constructs, determined by western blot analysis with antisera reactive with SARS-CoV S, was evaluated after transfection of the indicated plasmid expression vectors in 293T cells. Arrows indicate specific Δ CD (upper) and Δ TM (lower) bands.

assessed using an enzyme-linked immunosorbent assay (ELISA) with lectin-captured Δ TM protein expressed in 293T cells (see Methods). Substantial end-point dilution antibody titres were observed in all groups, ranging from $\sim 1:400$ to $\sim 1:2,000$ (Fig. 2b, left panel).

To analyse the ability of these S glycoprotein plasmids to elicit a neutralizing antibody response, a pseudotyping assay was performed¹⁵. Animals immunized with the SARS-CoV Δ CD expression vector induced substantial neutralizing antibody titres ranging from 1:50 to 1:150, unlike the vector control (Fig. 2b, middle panel). The S expression vector with the transmembrane domain deleted, Δ TM, also induced neutralizing antibodies, but the titres were lower (from 1:25 to 1:75). The vector with the partial cytoplasmic domain deletion induced the optimal response (Fig. 2b, right panel). Similar neutralization was observed when the complete

cytoplasmic domain was removed (data not shown), suggesting that synthesis of the glycoprotein on the cell surface, without the cytoplasmic domain, is required for this response. This may be because it gives rise to a more native structure relevant to the function of this virus. Possibly, the transmembrane region helps to form a more physiological form of the S protein by anchoring the protein on the membrane, preserving conformational determinants and/or stabilizing the formation of the putative trimer. The same antisera were also used in a SARS-CoV microneutralization assay to assess the neutralization titre against SARS-CoV. Again, antisera from Δ CD-vaccinated mice elicited the most potent neutralizing antibody responses (Fig. 2b, right panel).

Infection of non-human primates with SARS-CoV has been reported, with evidence of pulmonary pathology and seroconversion¹⁶. More recently, an animal model that may better reproduce the signs and time course of the human disease has been described. In this model, which uses ferrets, the development of pneumonitis and increased viral replication have been observed¹⁷. A murine model of SARS-CoV replication has also been described recently, in which intranasal administration of 10^4 50% tissue culture infectious dose (TCID₅₀) units of SARS-CoV leads to virus replication in the lungs and nasal turbinates within 1–2 days¹⁸. This model afforded us the opportunity to examine immune protection against viral replication in the respiratory tract as a measure of vaccine efficacy. BALB/c mice were immunized with the various plasmid DNAs encoding S protein and challenged 30 days after the final boost. Animals were challenged intranasally with 10^4 TCID₅₀ units of SARS-CoV (Urbani strain), and viral replication in the respiratory tract was measured 2 days later. In this analysis, the most potent immunogen, SARS Δ CD, led to $>10^6$ -fold reduction in viral load in the lungs compared with a control group injected with vector alone, in which mean viral titres of $>10^8$ were observed (Fig. 3a; $P = 0.0079$). Although the S plasmid with the deleted transmembrane domain elicited a lower neutralizing antibody response, it was equally effective in reducing viral replication in the lungs. A 60- to 300-fold reduction of virus titre in the nasal turbinates was also observed (Fig. 3b; $P = 0.0079$). In both cases, evidence of pro-

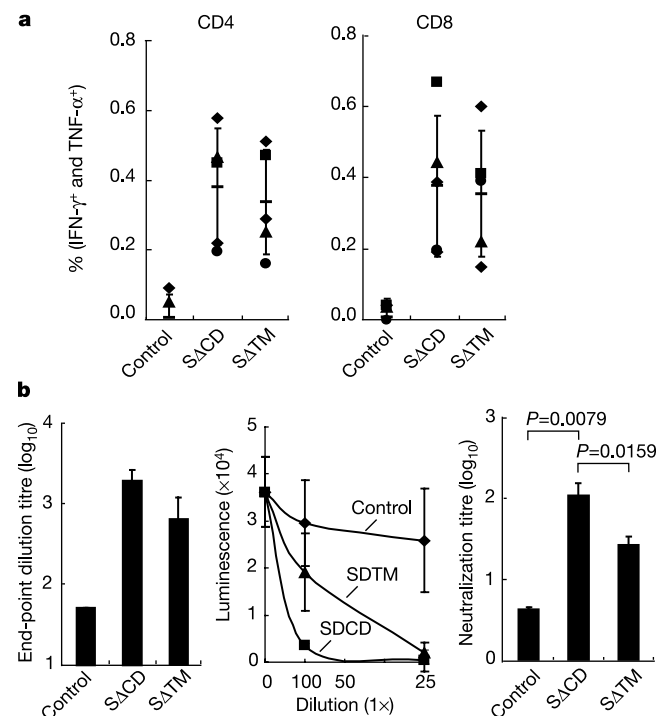


Figure 2 Immune responses to SARS-CoV DNA vaccination in BALB/c mice.

a, Intracellular cytokine staining was performed to quantify the percentage of activated T cells that produce either IFN- γ or TNF- α in response to stimulation with overlapping S peptide pools in CD4 (left) or CD8 (right) lymphocytes from mice ($n = 5$ per group) immunized with empty plasmid vector (control) or mice ($n = 5$ per group) immunized with the indicated plasmid at weeks 0, 3 and 6. Immune responses were measured 10 days after the final boost. Non-stimulated cells gave responses similar to those of the control subjects, at background levels. Symbols indicate the response of each individual animal, and the median value is shown (horizontal bar). **b**, Antibody responses induced by plasmid DNA vaccination against the SARS-CoV S protein. End-point dilution ELISA titres of SARS-CoV S-specific antibodies (left panel) in serum of vaccinated animals collected 10 days after the final boost were determined by optical density as described in the Methods. Neutralization by antisera from mice immunized with the relevant SARS-CoV S mutant or no insert (control) plasmid DNA vectors at the indicated concentrations was measured using the luciferase assay with S pseudotyped lentiviral vectors (middle panel). Reduction of gene transfer was observed with immune sera in a dose-dependent fashion. Twofold dilutions of heat-inactivated sera were tested in a microneutralization assay for the presence of antibodies that neutralized the infectivity of 100 TCID₅₀ of SARS-CoV in Vero cell monolayers, using four wells per dilution on a 96-well plate (right panel). The presence of viral cytopathic effect (cpe) was read on days 3 and 4. The dilution of serum that completely prevented cpe in 50% of the wells was calculated by the Reed Muench formula. Data are presented as the mean $\pm$ s.e. for each group. A non-parametric two-tailed t -test (Mann–Whitney) was used for statistical analysis, and the relevant P values are indicated (**b**, right panel).

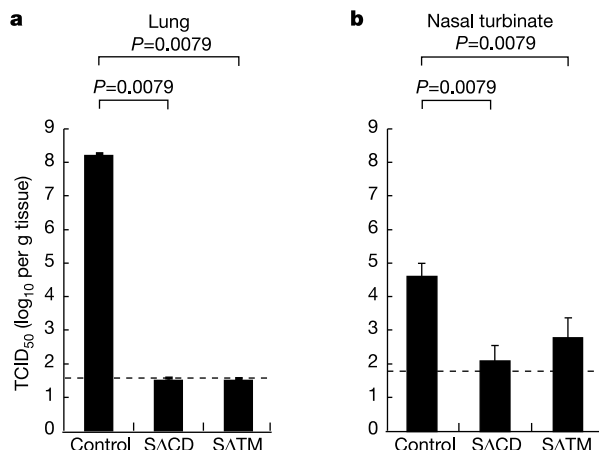


Figure 3 Protection against pulmonary SARS-CoV replication after challenge in BALB/c mice following DNA vaccination. Immunization and challenge were performed in mice as described previously¹⁸, and viral replication (mean log₁₀TCID₅₀ per g tissue with standard error) in the lower (**a**) and upper (**b**) respiratory tract after challenge with SARS-CoV was measured for five immunized animals inoculated with Δ CD, Δ TM or empty plasmid vector control. The lower limit for detection of SARS-CoV replication was 1.5 TCID₅₀ per g in the lungs and 1.8 TCID₅₀ per g in the nasal turbinates. A non-parametric two-tailed t -test (Mann–Whitney) was used for statistical analysis. Log-transformed virus titres were compared, and statistical significance was assigned to the differences, both with a P value of 0.0079.

ductive SARS-CoV replication was not observed in mice vaccinated with plasmid DNAs encoding Δ CD or Δ TM.

To define the mechanism of protection, T-cell depletion with specific monoclonal antibodies was performed, and depletion in the lung and spleen was confirmed (Supplementary Fig. 1). Depletion with CD4 or CD8, alone or in combination with CD90, did not affect vaccine-induced immunity (Fig. 4a). This finding was con-

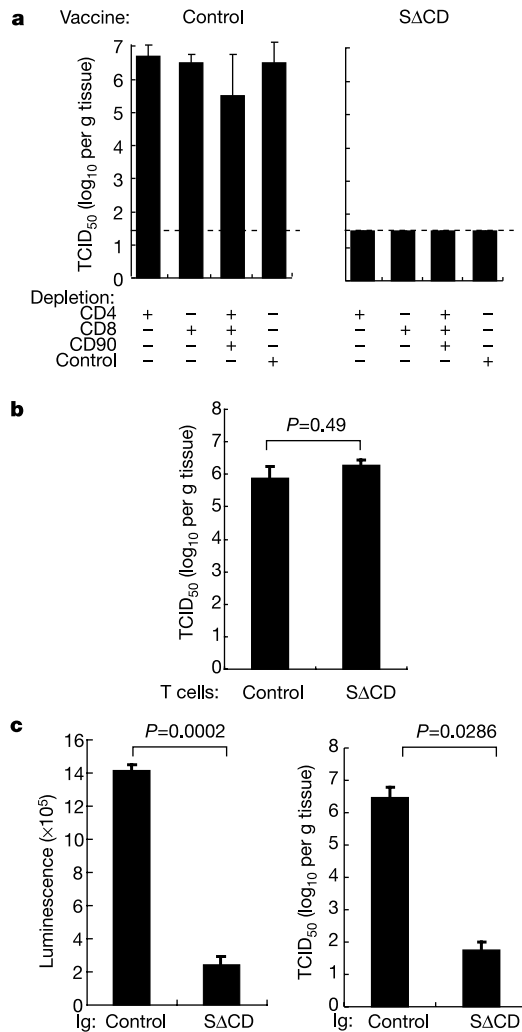


Figure 4 Immune mechanism of protection: T-cell depletion, adoptive transfer and antibody passive transfer. **a**, Monoclonal rat anti-mouse anti-CD4, CD8 or CD4/CD8/CD90 were used to deplete T cells in Δ CD and control vaccinated mice ($n = 4$) as described previously²³. Mice were then challenged with SARS-CoV 48 h later. Viral replication in the lungs was measured as described in Fig. 3. Data represent mean $\pm$ s.e., and no statistically significant difference was observed between groups depleted as shown. **b**, Lack of protection against SARS-CoV replication in the lungs after adoptive T-cell transfer from vaccinated mice. Each naive mouse ($n = 4$) received 3×10^7 purified T cells from a donor mouse confirmed to respond to the vaccine (immune) or from a non-immune mouse (control). The recipient mice were challenged 24 h after adoptive T-cell transfer. Viral titre in lungs was measured as described in Fig. 3. There was no statistically significant difference between these two groups using the non-parametric Mann-Whitney two-tailed t -test ($P = 0.49$). **c**, Protection against SARS-CoV replication in the lungs after passive transfer of immune IgG from vaccinated mice. Purified IgG from Δ CD or control vaccinated donor mice ($n = 4$) was passively transferred into recipient naive mice ($n = 4$). Serum from recipient mice was collected one day before challenge, and neutralization was confirmed using the S pseudotyped lentiviral vector (left panel). Recipient mice ($n = 4$ per group) were challenged 24 h after IgG transfer with 10^4 TCID₅₀ SARS-CoV and SARS-CoV replication was measured as described in Fig. 3. The non-parametric two-tailed t -test (Mann-Whitney) revealed a statistically significant difference of $P = 0.0286$ between these two groups (right panel).

firmed when the role of T-cell immunity was further analysed by adoptive T-cell transfer; donor immune T cells were unable to reduce pulmonary viral replication in recipient animals (Fig. 4b; $P = 0.49$). By contrast, passive transfer of purified IgG from immunized, but not control, mice provided immune protection (Fig. 4c, right panel; $P = 0.0286$) comparable to that observed in DNA-vaccinated animals (Fig. 3a). These findings indicate that immune T cells do not control pulmonary virus replication in this animal model, although it remains possible that T cells contribute to viral clearance if replication persists.

In this report, DNA vaccination has been used to induce cellular and humoral immunity to the SARS-CoV S glycoprotein. The humoral immune response includes the generation of neutralizing antibodies. This humoral immunity alone can inhibit pulmonary viral replication in a murine challenge model and suggests that DNA vaccination with the SARS-CoV S glycoprotein gene results in protective immunity. The SARS-CoV is a novel coronavirus, but vaccines for other human coronaviruses have not been successfully developed¹⁴. But considerable experience in developing vaccines against common veterinary coronaviruses has been obtained with animal coronaviruses^{11,12}. The feline infectious peritonitis virus is a coronavirus that causes significant morbidity and mortality in domestic animals for which vaccines have been investigated, but the development of these vaccines has been complicated by possible immune potentiation of the disease and by viral evolution¹⁹. In the present study, T-cell depletion and immune IgG passive transfer were used to assess the importance of humoral responses in protection against SARS-CoV challenge. These results suggest that antibodies against SARS-CoV S glycoprotein protect against SARS-CoV challenge and do not enhance infection in this animal model. The cellular immune response appears to play a limited role in protection, although studies of animal coronaviruses have suggested that both cellular and humoral immunity contribute to protection during persistent infection^{13,14}.

DNA vaccines have been used in a variety of animal models for different infectious diseases, including influenza, HIV, Ebola, West Nile and other viruses (reviewed in ref. 20). Despite their success in animal models, this approach has only recently been used in human studies, and its potential to protect against human diseases has yet to be established. For this reason, as well as the limitations of the animal model for SARS-CoV infections, it is important to ascertain the immunogenicity of these plasmids in humans. Should the response be suboptimal, it can be readily augmented using prime-boost combinations with inactivated viral vaccine candidates or with adenoviral or poxvirus vectors. The definition of effective viral genes reported here can guide the choice of inserts for such gene-based vaccination approaches. For example, the Δ CD mutant can be expressed in other vector delivery systems for analysis, alone or in various combinations. Although the murine model allows assessment of various antiviral approaches and examination of the effect of immunity on pulmonary virus replication, it does not show sustained infection or high lethality. The model of infection that has been described in ferrets shows prominent pulmonary pathology associated with infection, and the efficacy of such vaccines in this and other pathogenic animal models will be of interest in future studies. At the same time, it is not known whether even the ferret model faithfully replicates the complex pathology and symptomatology of the human disease, and it will be necessary to test potential SARS vaccine candidates for their immunogenicity, safety and efficacy in humans in the event of future outbreaks. □

Methods

Immunogen and plasmid construction

Plasmids encoding different versions of SARS-CoV spike (S) protein (Urbani strain, GenBank AY278741) were synthesized using human-preferred codons as previously described¹⁵. Protein expression was confirmed by western blot analysis²¹ with serum from a recovered patient (provided by W. Bellini).

Vaccination

Female BALB/c mice (6–8 weeks old; Charles River Labs) were immunized with 25 µg of plasmid DNA in 200 µl of PBS (pH 7.4) at weeks 0, 3 and 6.

Flow cytometric analysis of intracellular cytokines

CD4⁺ and CD8⁺ T-cell responses were evaluated by using intracellular cytokine flow cytometry (ICC) for IFN-γ and TNF-α as previously described²¹ with peptide pools (17–19 mers overlapping by 10 amino acids, 2.5 µg ml⁻¹ each) covering the SARS-CoV spike protein. Cells were then fixed, permeabilized, and stained using rat monoclonal anti-mouse CD3, CD4, CD8, IFN-γ and TNF-α (BD-Pharmingen). The IFN-γ- and TNF-α-positive cells in the CD4⁺ and CD8⁺ cell populations were analysed with the program FlowJo (Tree Star, Inc.).

ELISA for mouse anti-SARS-S IgG

The mouse anti-SARS-S IgG ELISA titre was measured using a modified lectin-capture method described previously²¹ except the Myc-tagged, transmembrane-domain-truncated SARS-CoV S protein (SARS-SATM-Myc) was used for capture.

Inhibition of viral gene transfer to measure mouse antibody titre

SARS-CoV spike pseudotyped lentiviruses expressing a luciferase reporter gene were produced by transfecting 293T cells with the following plasmids: 7 µg of pCMVΔR8.2, 7 µg of pHR⁺CMV-Luc and 400 ng CMV/R-SARS-S. Cells were transfected overnight, washed and replenished with fresh media. Forty-eight hours later, supernatants were harvested, filtered through a 0.45 µm syringe filter, aliquotted and used immediately or frozen at -80 °C. p24 levels were measured from different viral stocks using the Coulter HIV-1 p24 Antigen Assay kit (Beckman Coulter). Antisera were mixed with 100 µl of pseudoviruses at various dilutions and added to Vero cells in 48-well dishes (30,000 cells per well). Plates were washed and fresh media were added 14–16 h later. Forty-eight hours after infection, cells were lysed in mammalian cell lysis buffer (Promega). A standard quantity of cell lysate was used in a luciferase assay with luciferase assay reagent (Promega) according to the manufacturer's protocol.

Neutralization of SARS-CoV by mouse immune antisera

Twofold dilutions of heat-inactivated sera were tested in a microneutralization assay for the presence of antibodies that neutralized the infectivity of 100 TCID₅₀ of SARS-CoV on Vero cell monolayers, using four wells per dilution on a 96-well plate. The presence of viral cytopathic effect (cpe) was tested for on days 3 and 4. The dilution of serum that completely prevented cpe in 50% of the wells was calculated by the Reed Muench formula²².

Challenge of immunized mice with SARS-CoV

Vaccinated mice were lightly anaesthetized with isoflurane and inoculated with 50 µl of diluted virus (10⁴ TCID₅₀ of SARS-CoV; Urbani strain) intranasally according to institutional animal care and use guidelines in an ABSL3 facility. On day 2, mice were euthanized and lungs and nasal turbinates were removed and stored at -70 °C until the end of the study. The frozen tissues were thawed and homogenized in a 10% (lungs) or 5% (nasal turbinates) w/v suspension in Leibovitz 15 medium (Invitrogen), and virus titres were determined in Vero cell monolayers in 24- and 96-well plates. Virus titres are expressed as TCID₅₀ per g of tissue. The lower limit of detection of infectious virus in 10% and 5% w/v suspensions is 1.5 for lung and 1.8 log₁₀ TCID₅₀ per g for nasal turbinate homogenates, respectively.

Depletion of T-cell subsets *in vivo*

To deplete specific T-cell subsets, known rat monoclonal antibodies (anti-mouse CD4 (GK1.5), anti-mouse CD8(2.43) or anti-mouse CD90(30-H12)), prepared as described previously²³ and provided by S. L. Epstein (FDA), were administered intraperitoneally (1 mg each in 1 ml PBS) 48 h before challenge. The depletion was confirmed before challenge (Supplementary Fig. 1).

Passive transfer of immunoglobulin

IgG from mice immunized with plasmid DNA encoding ΔCD (immune) or no insert (controls) was purified from sera using a Protein A Antibody Purification Kit (Sigma), and neutralization activity was confirmed using the mouse pseudotyping assay. Briefly, 0.3 ml of purified IgG (from approximately 1 ml of serum) was administered intravenously into each recipient naive mouse (*n* = 4 per group) by tail vein injection 24 h before challenge. Sera were collected from recipient BALB/c mice 3 h post transfer to confirm the neutralizing antibody titre in recipient animals.

Adoptive T-cell transfer studies

T cells from vaccinated (immune) or nonimmune (control) mice were enriched using a Pan T-Cell Isolation Kit (Miltenyi Biotec). Approximately 3 × 10⁷ T cells in 0.5 ml PBS were administered into each recipient naive BALB/c mouse intravenously through the tail vein 24 h before challenge. There were four recipient mice per group.

Received 9 February; accepted 8 March 2004; doi:10.1038/nature02463.

1. Ksiazek, T. G. *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1953–1966 (2003).
2. Drosten, C. *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1967–1976 (2003).
3. Tsang, K. W. *et al.* A cluster of cases of severe acute respiratory syndrome in Hong Kong. *N. Engl. J. Med.* **348**, 1977–1985 (2003).
4. Lee, N. *et al.* A major outbreak of severe acute respiratory syndrome in Hong Kong. *N. Engl. J. Med.* **348**, 1986–1994 (2003).

5. Poutanen, S. M. *et al.* Identification of severe acute respiratory syndrome in Canada. *N. Engl. J. Med.* **348**, 1995–2005 (2003).
6. World Health Organization. Consensus document on the epidemiology of severe acute respiratory syndrome (SARS). *WHO* (<http://www.who.int/csr/sars/en/WHOconsensus.pdf>) (2003).
7. Rota, P. A. *et al.* Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**, 1394–1399 (2003).
8. Marra, M. A. *et al.* The genome sequence of the SARS-associated coronavirus. *Science* **300**, 1399–1404 (2003).
9. Snijder, E. J. *et al.* Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**, 991–1004 (2003).
10. Lai, M. M. C. & Holmes, K. V. in *Fields Virology* (ed. Knipe, D. M.) 1163–1185 (Lippincott, Williams & Wilkins, Philadelphia, 2001).
11. Saif, L. J. Coronavirus immunogens. *Vet. Microbiol.* **37**, 285–297 (1993).
12. Park, S. *et al.* Immune response of sows vaccinated with attenuated transmissible gastroenteritis virus (TGEV) and recombinant TGEV spike protein vaccines and protection of their suckling pigs against virulent TGEV challenge exposure. *Am. J. Vet. Res.* **59**, 1002–1008 (1998).
13. Sestak, K. *et al.* Active immunity and T-cell populations in pigs intraperitoneally inoculated with baculovirus-expressed transmissible gastroenteritis virus structural proteins. *Vet. Immunol. Immunopathol.* **70**, 203–221 (1999).
14. Holmes, K. V. in *Fields Virology* (ed. Knipe, D. M.) 1187–1203 (Lippincott Williams & Wilkins, Philadelphia, 2001).
15. Yang, Z.-Y. *et al.* pH-dependent entry of SARS coronavirus is mediated by the Spike glycoprotein and enhanced by dendritic cell-mediated transfer through DC-SIGN. *J. Virol.* (in the press).
16. Fouchier, R. A. *et al.* Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* **423**, 240 (2003).
17. Martina, B. E. *et al.* Virology: SARS virus infection of cats and ferrets. *Nature* **425**, 915 (2003).
18. Subbarao, K. *et al.* Prior infection and passive transfer of neutralizing antibody prevent replication of SARS coronavirus in the respiratory tract of mice. *J. Virol.* **78**, 3572–3577 (2004).
19. Perlman, S. Pathogenesis of coronavirus-induced infections. Review of pathological and immunological aspects. *Adv. Exp. Med. Biol.* **440**, 503–513 (1998).
20. Gurunathan, S., Klinman, D. M. & Seder, R. A. DNA vaccines: immunology, application, and optimization. *Annu. Rev. Immunol.* **18**, 927–974 (2000).
21. Kong, W. P. *et al.* Immunogenicity of multiple gene and clade human immunodeficiency virus type 1 DNA vaccines. *J. Virol.* **77**, 12764–12772 (2003).
22. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27**, 493–497 (1938).
23. Epstein, S. L. *et al.* Vaccination with DNA encoding internal proteins of influenza virus does not require CD8(+) cytotoxic T lymphocytes: either CD4(+) or CD8(+) T cells can promote survival and recovery after challenge. *Int. Immunol.* **12**, 91–101 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank J. McAuliffe and L. Vogel from the Laboratory of Infectious Diseases, M. St Claire, T. Tobery and staff at Bioqual for assistance in challenge studies, S. Rao and his staff for assistance in passive and adoptive transfer studies, J. Mascola for discussion, S. Epstein for supplying the monoclonal antibodies, A. Tislerics and T. Suhana for manuscript preparation, K. Stroud for preparation of figures, and members of the Nabel laboratory for advice and discussions.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.J.N. (gnabel@nih.gov).

Differentiating germ cells can revert into functional stem cells in *Drosophila melanogaster* ovaries

Toshie Kai & Allan Spradling

Howard Hughes Medical Institute Research Laboratories, Department of Embryology, Carnegie Institution of Washington, 115 W. University Parkway, Baltimore, Maryland 21210, USA

Many tissues including blood, skin, gut and germ cells are continuously maintained by tissue stem cells^{1–2}. Under certain conditions, however, other organs can undergo repair using stem-cell-like progenitors generated by cell de-differentiation³. Cell fates have been broadened experimentally^{4–7}, but mechanisms allowing de-differentiation to a stem cell state are poorly known. Germline stem cells begin to differentiate by forming interconnected germ cell cysts (cystocytes), and under certain

conditions male mouse cystocytes have been postulated to revert into functional progenitors^{8,9}. Here we report that four- and eight-cell *Drosophila* germline cystocytes generated either in second instar larval ovaries or in adults over-producing the BMP4-like stem cell signal Decapentaplegic efficiently convert into single stem-like cells. These de-differentiated cells can develop into functional germline stem cells and support normal fertility. Our results show that cystocytes represent a relatively abundant source of regenerative precursors that might help replenish germ cells after depletion by genotoxic chemicals, radiation or normal ageing. More generally, *Drosophila* cystocytes now provide a system for studying de-differentiation and its potential as a source of functional stem cells.

Drosophila melanogaster female germline stem cells (GSCs) constitute a highly favourable system for studying de-differentiation as a source of progenitor cells. GSCs reside in a specific microenvironment, a niche (see Fig. 1a), where they are attached tightly to stromal cap cells¹⁰ and maintained by *decapentaplegic* (*dpp*) signalling¹¹. After division, one daughter cell stays in the niche as a stem cell,

while the other moves out of the niche, de-represses *bag of marbles* (*bam*) expression¹², and begins a programme of cyst differentiation that ultimately produces one oocyte connected to 15 nurse cells by actin-rich ring canals¹³. Once initiated, cyst differentiation entails four divisions with incomplete cytokinesis, the stabilization of ring canals and the progressive, asymmetric growth of a complex cytoskeletal organelle known as the fusome (Fig. 1b)¹⁴. Cyst growth seems to be irreversible, as *Drosophila* germ cells more mature than cystoblasts have never been reported to de-differentiate.

One possible reason that cyst de-differentiation has not been observed is that cysts susceptible to reversion might never experience the appropriate conditions to elucidate this response. For example, Dpp reception in adult ovarioles is limited to the GSC niche (Fig. 1a, dark pink cells), not the more posterior region containing cysts¹⁵ (Fig. 1a, light pink cells). To look for a more favourable environment, we investigated the organization and *dpp*-responsiveness of germ cells in second instar (L2) larval ovaries (Fig. 1c). Before the late third instar when germ cells begin to associate with cap cells¹⁶, they are all thought to be equivalent, potential stem cell progenitors. *Bam* expression and cyst initiation only get under way at the larval-pupal transition¹⁶. We found that germ cell progenitors divide in a similar manner to GSCs¹⁴ (Fig. 1d, e). Their round fusomes partition asymmetrically while daughter cells remain transiently interconnected in a cell pair after telophase (Fig. 1e, inset). These short-lived germ cell pairs break apart when the ring canal pinches shut part way through the next cell cycle. Unlike the adult ovariole, however, Dpp signalling is activated throughout the entire L2 germ cell compartment, as indicated by anti-phospho-Mad (pMad) staining (Fig. 1f). The widespread nature of Dpp signalling in L2 larval ovaries suggests a possible favourable environment for cyst reversion.

Forcing the expression of Bam protein in GSCs causes them to differentiate into cysts¹⁷. We tested the stability of cysts in larval ovaries by subjecting larval germ cells to a brief dose of Bam protein using a heat-shock Bam transgene (*hs-bam*) (Fig. 2a). At 20 h after heat-shock treatment, administered at the start of the second larval instar, no single germ cells or cell pairs remained (Fig. 2b, compare left and right panels). Instead, only 4- and 8-cell cysts were found in numbers that corresponded closely to the starting number of single and transiently interconnected germ cell pairs, respectively (Table 1; see also Supplementary Information). The characteristic shape of the fusomes and appearance of ring canals in these larval cysts (Fig. 3a, b) were indistinguishable from adult cysts¹⁴. Over the next 20–30 h of larval development, however, the cysts became unstable

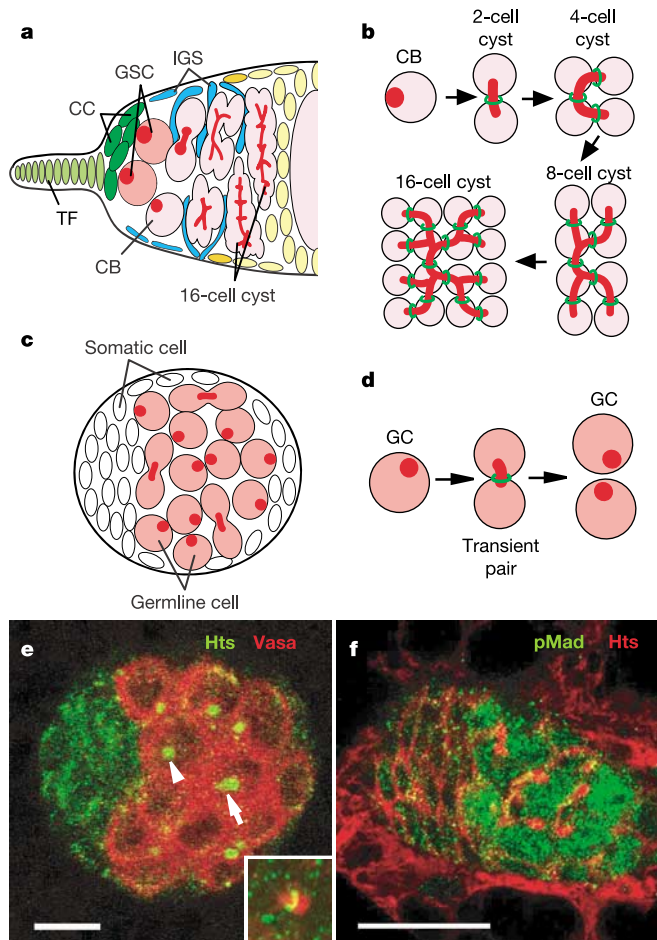


Figure 1 Adult ovariole and L2 larval ovary. **a**, Adult ovariole tip. Anterior is to the left. GSCs (dark pink; active Dpp signalling), cystoblasts (CB) and cyst cells (light pink; low Dpp signalling) are shown. GSCs reside in a niche composed of cap cells (CC; dark green) and terminal filament cells (TF; light green). Fusomes are coloured red. **b**, Cyst differentiation. Ring canals are coloured green. **c**, L2 larval ovary. Primordial germ cells (GC) are coloured dark pink, whereas somatic cells are white. **d**, Asymmetric germ cell cycle. Germ cell progenitors divide asymmetrically like GSCs¹⁴. **e**, L2 larval ovary showing Vasa (red) and Hts (green) staining. Compare with panel **c**. The inset is of a transient germ cell pair showing temporary ring canal (anti-Mucin D, yellow) and asymmetric fusome (anti-Hts, red). **f**, L2 larval ovary showing Hts (red) and phospho-Mad (green) staining. Scale bars, 10 μ m.

Table 1 Experimentally induced cysts in L2 ovaries

Ovary number	Single cells	Pair or two-cell cysts	Four-cell cysts	Eight-cell cysts
1	0	0	13	0
2	0	0	7	3
3	0	0	6	3
4	0	0	8	2
5	0	0	6	2
6	0	0	7	0
7	0	0	6	0
8	0	0	3	3
Average	0	0	7 $\pm$ 2.8	1.6 $\pm$ 1.4

Data were collected 20 h after heat-shock treatment.

Table 2 Apoptosis in treated and untreated larval ovaries

Cell type	Heat shock (number of apoptotic cells)	Control (number of apoptotic cells)
Germ line (30 h)	0	0
Somatic (30 h)	1.4 $\pm$ 1.7	1.5 $\pm$ 1.3
Germ line (40 h)	0	0
Somatic (40 h)	2 $\pm$ 1.6	3 $\pm$ 1.6

Time in the first column refers to time after heat-shock treatment.

and began to break down (Fig. 2c). By 50 h after heat shock, only single cells were found in heat-shocked ovaries (Fig. 2d, compare left and right panels).

The single germ cells that re-formed in the heat-treated ovaries must derive from cyst breakdown. First, the number of single germ cells 50 h after heat shock corresponded closely to the number of cystocytes after 20 h (Fig. 2f). To rule out the possibility that most of the cells die and only a few rare cells give rise to the surviving single germ cells, we counted the number of dying germ cells throughout the critical period. No apoptotic germline cells were observed at any time point (Table 2), although our method easily detected low levels of somatic cell apoptosis that occurred in both treated and control ovaries (Fig. 2e). Finally, visualizing dividing cells using anti-phospho-histone H3 confirmed that cell division in the heat-shocked ovaries was very low (Fig. 2g). Consequently, all of the germ cells at 50 h after heat shock must have derived from the 4- and 8-cell cysts present at 20 h after heat shock.

If cysts are truly breaking down, then we should observe specific intermediates during this process. Consequently, we stained ovaries at various times after heat shock with anti-Hts to label fusomes and anti-Anillin to label ring canals and cell nuclei¹⁴. At 20 h after heat shock, the cysts contained normally branched fusomes and ring canals (Fig. 3a, b); however, at 30 h, abnormalities in fusome structure were apparent. Most fusomes remained intact, but were much thinner near ring canals, which seemed to be constricting (Fig. 3c). At 40 h after heat shock, many ring canals were completely closed, pinching the fusome into pieces (Fig. 3d). These observations indicate that cysts break into single cells by ring canal

closure and fusome fragmentation. At 50 h after heat shock, ring canal remnants could be seen clearly in the single germ cells (Fig. 3e), further confirming their origin.

To determine whether the reverted cystocytes could form functional stem cells we studied the adult females that developed from heat-shocked larvae. Ovariole formation is controlled by somatic cells¹⁸, independently of whether there are stem cells available to populate them. We found that adult females that received the heat-shock treatment as larvae contained the same number of ovarioles with similar numbers of GSCs as from untreated animals (Table 3). Therefore, during ovary formation reverted cystocytes must populate the forming niches like normal stem cell precursors. Moreover, the environment surrounding the stem cells was indistinguishable from controls and constituted a normal niche, as indicated by anti-pMad staining (Fig. 3f). Five-day-old treated and control females laid normal eggs at about the same rate, and about 90% of eggs from both groups hatched and developed to adulthood. The re-constituted stem cells appeared to have a normal half-life, as 25-day-old treated flies laid a similar number of eggs as age-matched controls. These results indicate that GSCs derived from reverted cyst cells are equivalent to normal GSCs.

To determine whether cyst de-differentiation can also be induced in adults, we first generated GSC tumours in adult flies carrying a *hs-bam* transgene by driving *dpp* expression in somatic ovarian cells¹⁵ (Fig. 4a). When Bam is expressed after a heat shock in these flies, the multiple GSCs in each ovariole immediately and synchronously begin to differentiate and soon generate 4-cell and 8-cell cysts (Fig. 4b). There were no single cells or cell pairs remaining at

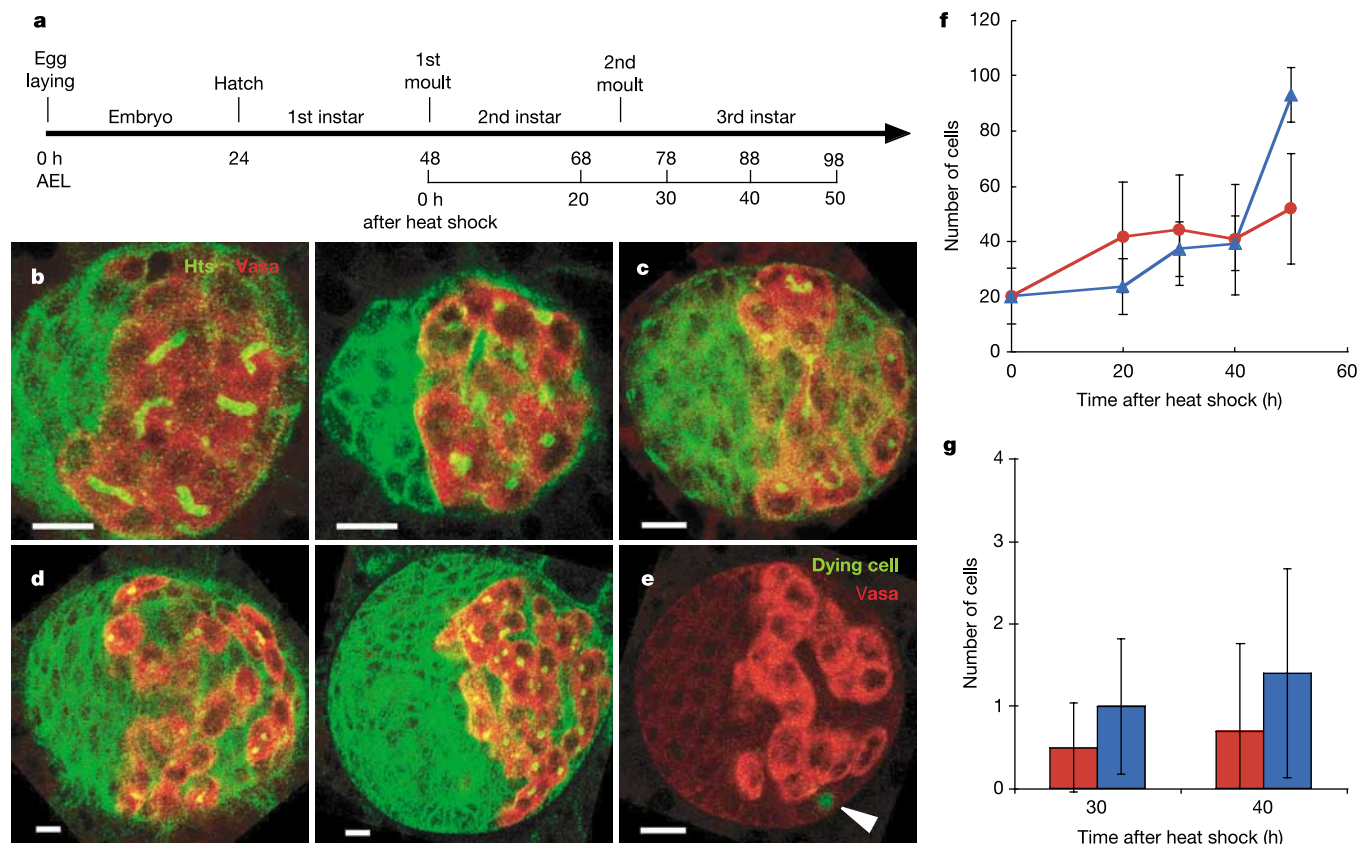


Figure 2 Germ cell cysts in L2 larval ovaries revert to single cells. **a**, Larvae 48 h after egg laying (AEL) were heat-shocked, and ovaries analysed at various times. **b–d**, Hts (fusome, green) and Vasa (red) staining is shown. **b**, Bam-treated ovary 20 h after heat shock containing 4- and 8-cell cysts (left). The right panel shows a 68 h control ovary. **c**, Bam-treated ovary 30 h after heat shock; note that the fusome structure appears abnormal. **d**, Bam-treated ovary 50 h after heat shock containing single germ cells or transient cells

pairs (left); compare with control 98 h ovaries (right). **e**, Apoptotic somatic cell (green) in Bam-treated ovary 30 h after heat shock. Vasa staining is coloured red. Scale bars, 10 μ m. **f**, Germ cell number per ovary. The red line indicates ovaries treated with Bam, whereas the blue line indicates control ovaries. **g**, Phospho-histone H3-positive cells. The red columns indicate ovaries treated with Bam, whereas the blue columns indicate control ovaries.

Table 3 Reverted stem cells function normally

Parameter	Heat-shocked	Control
GSCs (day 5)	2.1 ± 0.32	2.1 ± 0.56
Ovarioles (day 5)	19 ± 1.3	18.7 ± 1.3
Eggs laid per female per day (day 5)	37	34
Eggs laid per female per day (day 25)	21	22

20 h after heat shock; however, as observed in the L2 larval ovary, by 30 h after heat shock the cysts stopped growing and started to break down (Fig. 4c). By 40 h after heat shock closed ring canals and ring canal remnants were common (Fig. 4d). By 50 h after heat shock almost all of the germ cells were single (Fig. 4e). Thus, cysts break down in adult flies with the same kinetics and through the same process of ring canal closure and fusome scission as in larvae. The continued presence of excess Dpp prevented us from testing whether the reverted cells can function as stem cells or differentiate. Thus, cyst de-differentiation is clearly not limited to larval gonads but also occurs in adult flies, where it will be amenable to detailed molecular study.

Several factors potentially contribute to cyst instability in our experiments (Fig. 4f, g). First, reversion may be stimulated by Dpp signalling. Cysts in Bam-treated adult ovarioles do not receive Dpp signals¹⁵ (Fig. 4f, light pink cells), whereas all germ cells in Bam-treated early larval ovaries (data not shown) or adult tumours¹⁵ are Dpp-responsive (Fig. 4g, dark pink cells). Dpp signals are known to

repress *bam* expression¹². Our results suggest that these signals can stimulate ring canal closure and fusome scission, which is a normal process in cycling GSCs with active Dpp signalling¹⁴. The loss of normal cellular interactions with somatic inner germarium sheath (IGS) cells, which envelope growing adult cysts, may be another destabilizing factor (Fig. 4, blue cells). The nature of these interactions and whether they contribute to cyst stability will require further study.

The generation of new stem cells from cystocytes probably has a biological significance in several situations. Normally, cyst reversion must be prevented so that oogenesis can proceed productively. In addition to the downregulation of Dpp signalling and enhanced IGS cell interactions experienced by newly formed cysts, ongoing changes in ring canal structure¹⁹ (including the activity of associated Src kinases^{20,21}) may prevent breakdown. After depletion of germline stem cells by environmental toxins, chemotherapy, radiation exposure or normal ageing, however, cysts may be destabilized and generate replacement stem cells. As the number of cystocytes exceeds the number of GSCs by 1–3 orders of magnitude in many organisms¹³, including male mice⁹, this would represent a very large increase in the potential progenitor population. It is known that mice whose germ cells have been depleted can achieve a reconstituted germ cell population through the injection of male germ cells²². Regenerating germ cell colonies have been assumed to arise from input GSCs, but our results indicate that transferred cystocytes may break down and produce some colony-forming cells. Indeed, the

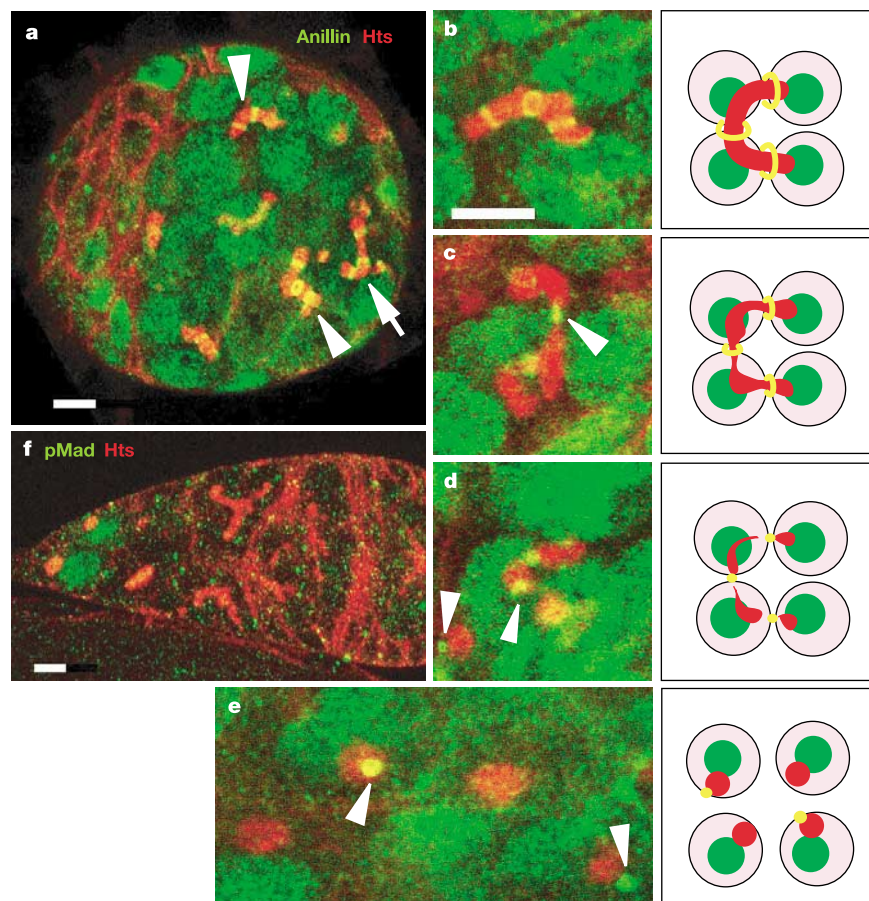


Figure 3 Reverting cystocytes close ring canals and become functional GSCs. **a–e**, Bam-treated larval ovaries. Hts (red) and Anillin (ring canals and nuclei, green) staining is shown. The right panels of **b–e** show corresponding illustrative diagrams, where the fusome (red), nuclei (green) and ring canals (yellow) are shown. **a**, Ovary 20 h after heat shock showing induced 4- (arrowhead) and 8-cell (arrow) cysts with normal morphology. **b**, Ovary 20 h after heat shock showing single 4-cell cyst. **c**, Ovary 30 h after

heat shock showing a 4-cell cyst with constricting ring canal (arrowhead). **d**, Ovary 40 h after heat shock showing closed ring canal and broken fusome (arrowhead). **e**, Ovary 50 h after heat shock showing separated cells with ring canal remnants (arrowheads). **f**, Morphologically normal adult ovariole from a Bam-treated larva containing normal GSCs with active Dpp signalling (green). Hts (red) and pMad (green) staining is indicated. Scale bars, 5 μm.

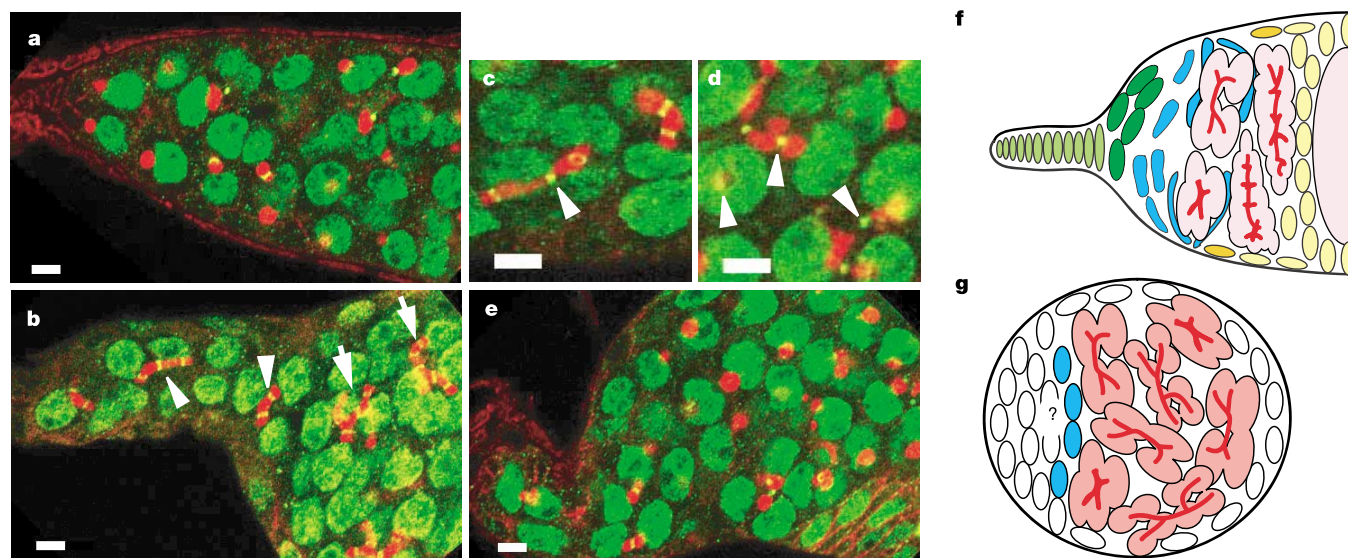


Figure 4 GSC reversion in adult flies. **a–e**, Untreated (**a**) and Bam-treated (**b–e**) adult ovarioles from flies overexpressing Dpp. **b**, Ovariole 20 h after heat shock showing induced 4- (arrowhead) and 8-cell (arrow) cysts with normal morphology. No single cells or cell pairs were detected ($n > 100$). **c**, Ovariole 30 h after heat shock showing a 4-cell cyst with a constricting ring canal (arrowhead). **d**, Ovariole 40 h after heat shock showing

closed ring canals and broken fusomes (arrowheads). **e**, Ovariole 50 h after heat shock showing almost complete cyst breakdown. Scale bars, 5 μ m. **f**, **g**, Schematic diagram of an adult ovariole 2 days after heat shock (**f**) and of a larval ovary 20 h after heat shock (**g**). Colours are the same as in Fig. 1a.

ability of mouse A-type spermatogonia (that is, cystocytes) at certain stages to regenerate earlier precursors has long been postulated^{8,9}.

In broader terms, our experiments show that stem cells that have embarked on a differentiation pathway and undergone marked morphological changes can still revert back to their original state. Little is currently known about the nuclear changes involved in the transition between a GSC and the various stages of cystocyte development; however, our experiments show that significant differences in cytoplasmic structure are not a barrier to stem cell reversion. Previous work with cellular intermediates in a number of pathways shows that developmental plasticity can be increased by altering transcriptional programmes^{4–7}. Our findings encourage a focus on differentiating cells, in addition to stem cells, as a source of progenitors for tissue repair. □

Methods

Drosophila stocks

Either *y ry* or OregonR were used as wild type. We used the following fly stocks in our study: *w¹¹¹⁸*; *P[w⁺,hsp70 bam⁺]^{11d}* (ref. 17) and *w¹¹¹⁸* *c587-GAL4*; *UAS-dpp* (ref. 15). Further information on symbols and genes can be found in FlyBase (<http://flybase.bio.indiana.edu/>).

Cyst induction using *hs-bam*

For analysis of L2 larval ovaries, eggs from *w¹¹¹⁸*; *P[w⁺,hsp70 bam⁺]^{11d}* adults were collected for 3 h in plastic vials, and cultured at 25 °C. A single 1.5 h heat shock at 37 °C was given to larvae at 48 $\pm$ 1.5 h after egg laying. Subsequently, ovaries were dissected from heat-shocked larvae, stained with antibodies and examined as described below. To study cyst reversion in adults we increased the number of GSCs by overexpressing Dpp specifically in the somatic cells of adult ovariole tips using the GAL4 driver line *c587*. Two-day-old *w¹¹¹⁸* *c587-GAL4*; *P[w⁺,hsp70 bam⁺]^{11d}*/*UAS-dpp* adults were subjected to a single 2 h heat shock at 37 °C. Ovarioles dissected at various times thereafter were stained as for L2 ovaries.

Apoptosis assay

Apoptotic cells in ovarian tissue were identified by staining with Apoptag as described previously¹⁵.

Immunostaining and fluorescence microscopy

Ovaries and adult ovarioles were stained and mounted as described¹⁵. Germ cells were identified by staining for the germ-cell-specific protein Vasa. Fusomes were visualized with antibodies to the fusome structural protein Hu-li tai shao (Hts). Ring canals were visualized with antibodies to the actin-binding protein Anillin¹⁴, which also labels cell nuclei, or with anti-Mucin D²³. Active Dpp signalling was detected by staining for the

phosphorylated form of the *Drosophila* Smad protein Mad (pMad)²⁴. We used the following anti-sera: anti-pMad (PS1) (gift from P. ten Dijke; 1:200); anti-Hts mouse monoclonal antibody 1B1 (Developmental Studies Hybridoma Bank; 1:50); rabbit or rat anti-*Drosophila* Vasa (gift from P. Lasko; 1:8,000 or 1:1,000, respectively); anti-Anillin (provided by C. Field; 1:200)¹⁴. Rabbit anti-*Drosophila* mucin D antiserum (1:1,000) was provided by A. A. Kramerov²³; rabbit anti-phospho-histone H3 was obtained from Upstate Biotechnology (1:100). Secondary antibodies were goat anti-mouse IgG or goat anti-rabbit IgG conjugated to either Alexa 488 or Alexa 568 (Molecular Probes; 1:400), and Cy5-conjugated anti-rat IgG (Jackson ImmunoResearch; 1:1,000). Images were taken using a Leica TCS NT or Leica TCS SP2 confocal microscope.

Received 19 December 2003; accepted 24 February 2004; doi:10.1038/nature02436.
Published online 14 March 2004.

1. Spradling, A. C., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414**, 98–104 (2001).
2. Watt, F. M. & Hogan, B. L. M. Out of Eden: stem cells and their niches. *Science* **287**, 1427–1430 (2000).
3. Stocum, D. L. Amphibian regeneration and stem cells. *Curr. Top. Microbiol. Immunol.* **280**, 1–70 (2000).
4. Mikkola, I., Heavey, B., Horcher, M. & Busslinger, M. Reversion of B cell commitment upon loss of Pax5 expression. *Science* **297**, 110–113 (2002).
5. Morrison, S. J. *et al.* Transient Notch activation initiates an irreversible switch from neurogenesis to gliogenesis by neural crest stem cells. *Cell* **101**, 499–510 (2000).
6. Prohaska, S. S., Scherer, D. C., Weissman, I. L. & Kondo, M. Developmental plasticity of lymphoid progenitors. *Semin. Immunol.* **14**, 377–384 (2002).
7. Pearson, B. J. & Doe, C. Q. Regulation of neuroblast competence in *Drosophila*. *Nature* **425**, 624–628 (2003).
8. Clermont, Y. & Bustos-Obregon, E. Re-examination of spermatogonial renewal in the rat by means of seminiferous tubules mounted 'in toto'. *Am. J. Anat.* **122**, 237–247 (1968).
9. De Rooij, D. G. & Russell, L. D. All you wanted to know about spermatogonia but were afraid to ask. *J. Androl.* **21**, 776–798 (2000).
10. Song, X., Zhu, C. H., Doan, C. & Xie, T. Germline stem cells anchored by adherens junctions in the *Drosophila* ovary niches. *Science* **296**, 1855–1857 (2002).
11. Xie, T. & Spradling, A. C. *decapentaplegic* is essential for the maintenance and division of germline stem cells in the *Drosophila* ovary. *Cell* **94**, 251–260 (1998).
12. Chen, D. & McKearin, D. Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. *Curr. Biol.* **13**, 1786–1791 (2003).
13. de Cuevas, M., Lilly, M. A. & Spradling, A. C. Germline cyst formation in *Drosophila*. *Annu. Rev. Genet.* **31**, 405–428 (1997).
14. de Cuevas, M. & Spradling, A. C. Morphogenesis of the *Drosophila* fusome and its implications for oocyte specification. *Development* **125**, 2781–2789 (1998).
15. Kai, T. & Spradling, A. C. An empty *Drosophila* stem cell niche reactivates the proliferation of ectopic cells. *Proc. Natl Acad. Sci. USA* **100**, 4633–4638 (2003).
16. Zhu, C. H. & Xie, T. Clonal expansion of stem cells during niche formation in *Drosophila*. *Development* **130**, 2579–2588 (2003).
17. Ohlstein, B. & McKearin, D. Ectopic expression of the *Drosophila* Bam protein eliminates oogenic germline stem cells. *Development* **124**, 3651–3662 (1997).
18. Godt, D. & Lasko, F. A. Mechanisms of cell rearrangement and cell recruitment in *Drosophila* ovary morphogenesis and the requirement of *bric-a-brac*. *Development* **121**, 173–187 (1995).

19. Robinson, D. N., Cant, K. & Cooley, L. Morphogenesis of *Drosophila* ovarian ring canals. *Development* **120**, 2015–2025 (1994).
20. Dodson, G. S., Guarnieri, D. J. & Simon, M. A. Src64 is required for ovarian ring canal morphogenesis during *Drosophila* oogenesis. *Development* **125**, 2883–2892 (1998).
21. Roullet, E. M., Panzer, S. & Beckendorf, S. K. The Tec29 tyrosine kinase is required during *Drosophila* embryogenesis and interacts with Src64 in ring canal development. *Mol. Cell* **1**, 819–829 (1998).
22. Brinster, R. L. Germline stem cell transplantation and transgenesis. *Science* **296**, 2174–2176 (2002).
23. Kramerova, I. A. & Kramerov, A. A. Mucinoprotein is a universal constituent of stable intercellular bridges in *Drosophila melanogaster* germ line and somatic cells. *Dev. Dyn.* **216**, 349–360 (1999).
24. Tanimoto, H., Itoh, S., ten Dijke, P. & Tabata, T. Hedgehog creates a gradient of DPP activity in *Drosophila* wing imaginal discs. *Mol. Cell* **5**, 59–71 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank R. Cox, B. Ohlstein and E. deCotto for comments on the manuscript.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.C.S. (spradling@ciwemb.edu).

AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus

Yasuhiko Minokoshi¹, Thierry Alquier¹, Noboru Furukawa¹, Young-Bum Kim¹, Anna Lee¹, Bingzhong Xue¹, James Mu², Fabienne Foulle³, Pascal Ferré³, Morris J. Birnbaum², Bettina J. Stuck¹ & Barbara B. Kahn¹

¹Division of Endocrinology, Diabetes and Metabolism, Beth Israel Deaconess Medical Center and Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215, USA

²Howard Hughes Medical Institute, The Cox Institute, University of Pennsylvania Medical School, Philadelphia, Pennsylvania 19104, USA

³Unit 465 INSERM, Centre de Recherches Biomedicales des Cordeliers, 75270 Paris Cedex 6, France

Obesity is an epidemic in Western society, and causes rapidly accelerating rates of type 2 diabetes and cardiovascular disease. The evolutionarily conserved serine/threonine kinase, AMP-activated protein kinase (AMPK), functions as a 'fuel gauge' to monitor cellular energy status¹. We investigated the potential role of AMPK in the hypothalamus in the regulation of food intake. Here we report that AMPK activity is inhibited in arcuate and paraventricular hypothalamus (PVH) by the anorexigenic hormone leptin, and in multiple hypothalamic regions by insulin, high glucose and refeeding. A melanocortin receptor agonist, a potent anorexigen², decreases AMPK activity in PVH, whereas agouti-related protein, an orexigen², increases AMPK activity. Melanocortin receptor signalling is required for leptin and refeeding effects on AMPK in PVH. Dominant negative AMPK expression in the hypothalamus is sufficient to reduce food intake and body weight, whereas constitutively active AMPK increases both. Alterations of hypothalamic AMPK activity augment changes in arcuate neuropeptide expression induced by fasting and feeding. Furthermore, inhibition of hypothalamic AMPK is necessary for leptin's effects on food intake and body weight, as constitutively active AMPK blocks these effects. Thus, hypothalamic AMPK plays a critical role in hormonal and nutrient-derived anorexigenic and orexigenic signals and in energy balance.

Multiple factors regulate food intake, including hormones, fuels

and behaviour. Leptin, insulin, and melanocortin receptor agonists are potent anorexigen². Absence of leptin³ or melanocortin receptors² or neuron-specific deletion of insulin receptors⁴ causes hyperphagia and obesity. Glucose⁵ and fatty acids⁶ also modulate the activity of neurons that regulate appetite. The central nervous system signalling networks that mediate the actions of hormones and nutrients on energy balance are of key interest. AMPK is a heterotrimer consisting of catalytic α -subunits and regulatory β - and γ -subunits¹. AMPK is regulated by the cellular AMP/ATP ratio and by upstream kinases^{1,7}. AMPK is activated by stress and regulates cellular metabolism by inhibiting energy consuming pathways and inducing pathways that generate ATP¹. It is widely expressed—locations include neurons and glial cells^{8,9}. We hypothesized that hypothalamic AMPK might mediate hormonal and nutrient effects on food intake and energy balance.

First we assessed whether leptin alters hypothalamic AMPK activity. The accuracy of dissection of hypothalamic nuclei was assessed by measuring the expression of neuropeptides expressed in distinct hypothalamic regions¹⁰ (Supplementary Methods and Supplementary Fig. 1). Intraperitoneal (i.p.) injection of leptin in free-moving fasted mice decreased α 2AMPK activity in PVH and arcuate hypothalamus (ARH) at 3 and 6 h after injection (Fig. 1a). In contrast, leptin did not affect AMPK activity in other medial hypothalamic regions (ventromedial hypothalamus, VMH, and dorsomedial hypothalamus, DMH) or in lateral hypothalamus (LH). Moreover, α 1AMPK activity did not change in response to leptin in any hypothalamic region (not shown). Baseline α 2AMPK activity was high in the cortex and was unaltered by leptin administration. Intrahypothalamic injection of leptin had similar effects to i.p. injection and inhibited α 2AMPK activity selectively in PVH and ARH (Fig. 1b). These effects could be direct or indirect via neuronal networks. In contrast to the restricted effect of leptin on AMPK activity, i.p. and intrahypothalamic leptin both increased STAT3 phosphorylation in all hypothalamic regions, with greater effects in ARH and VMH/DMH than in PVH and LH (Fig. 1c).

The melanocortin (MC) 4 receptor pathway potently inhibits appetite and at least partially mediates the anorexigenic effect of leptin². Injection of MT-II, an agonist for MC3 and 4 receptors², into the lateral ventricle (intracerebroventricularly, i.c.v.) inhibited α 2AMPK activity in PVH (Fig. 1d), similar to leptin. There was no effect of MT-II in the medial hypothalamus and, similar to leptin, there was no effect in LH or cortex. Insulin is also a potent anorexigenic hormone^{2,11}. Insulin i.c.v. reduced α 2AMPK activity by 25–40% in all hypothalamic regions but not in cortex (Fig. 2d). Thus, the effect of insulin was more widespread in the hypothalamus than that of leptin or MT-II.

Brain glucose concentrations regulate neuronal firing rate⁵ and a rise in glucose may suppress feeding behaviour and alter autonomic tone⁵. Glucose injection (i.p.) in awake mice suppressed α 2AMPK activity in all hypothalamic regions but not cortex (Fig. 1e). Because systemic injection of glucose causes hyperglycaemia (12.2 ± 1.3 mM in i.p. glucose-injected mice; 5.1 ± 0.4 mM in saline-injected mice), which stimulates insulin secretion (0.89 ± 0.14 ng ml⁻¹ in i.p. glucose-injected mice; 0.09 ± 0.01 ng ml⁻¹ in saline-injected mice) and alters levels of substrates such as fatty acids, we subsequently injected glucose i.c.v. (100 μ g) to determine the impact of changes in brain glucose concentrations *per se*. Glucose injection (i.c.v.) did not change serum glucose (5.7 ± 0.2 mM) or insulin (0.12 ± 0.002 ng ml⁻¹) levels, compared with the saline-injected controls (glucose: 5.3 ± 0.3 mM, insulin: 0.11 ± 0.02 ng ml⁻¹). However, α 2AMPK activity was suppressed in all hypothalamic regions, similar to the effects of i.p. glucose (Fig. 1e). Thus, the AMPK pathway appears to respond coordinately to several anorexigenic inputs. Neither i.p. nor i.c.v. glucose or insulin or MT-II affected α 1AMPK activity.

Refeeding rapidly reduced AMPK activity in all hypothalamic nuclei, demonstrating the physiological significance (Fig. 1f). This

effect persisted until at least 6 h of refeeding (not shown). In mice that were refed after an overnight fast, $\alpha 2$ AMPK activity decreased in all hypothalamic regions (Fig. 1f), possibly owing to refeeding-induced elevation in serum glucose (9.7 ± 0.4 versus 6.4 ± 0.2 mM in fasted) and/or serum insulin (1.16 ± 0.32 versus 0.16 ± 0.03 ng ml⁻¹ in fasted). In contrast, agouti-related protein (AGRP) i.c.v. increased $\alpha 2$ AMPK activity from the refeeding to the fasting level in PVH only. Neuropeptide Y (NPY) did not change $\alpha 2$ AMPK activity in any hypothalamic region. Neither AGRP nor NPY altered $\alpha 2$ AMPK activity in the fasted state (not shown), probably owing to AMPK activity already being high.

AGRP is a specific competitive antagonist of MC3 and MC4 receptors¹²; some effects of leptin in PVH are mediated by the MC4 receptor². The reciprocal effects of AGRP and MT-II on AMPK activity in PVH indicate that PVH AMPK activity is regulated by MC3/4 receptors. Therefore, we examined regulation of AMPK activity in PVH in MC4 receptor knockout mice. $\alpha 2$ AMPK activity in PVH did not decrease in response to i.p. leptin or refeeding in MC4 receptor knockout mice (Fig. 1g). Thus, the decrease of $\alpha 2$ AMPK activity in PVH in response to leptin or refeeding depends on MC4 receptor signalling.

To determine whether modulation of hypothalamic AMPK activity alters food intake and body weight, we expressed constitutively active (CA, H150R mutation in the $\gamma 1$ subunit of AMPK) (Supplementary Methods) and dominant negative (DN) $\alpha 1$ (D157A)¹³ and $\alpha 2$ (K45R) (Supplementary Methods) subunits of AMPK in the medial hypothalamus using recombinant adenoviruses. We expressed $\alpha 1$ and $\alpha 2$ DN-AMPK simultaneously for maximal effects on AMPK activity, because in the absence of $\alpha 2$ AMPK, $\alpha 1$ AMPK can upregulate¹⁴ and in preliminary

experiments, $\alpha 1$ AMPK activity increased when we expressed only $\alpha 2$ DN-AMPK. The DN effect results from competition of the catalytically inactive α -subunits with endogenous wild-type α -subunits, for binding to β - and γ -subunits¹³. Control mice were injected with an adenovirus with no insert ('Null'). The VMH was chosen as the injection site to mimic the widespread effects of many anorexigenic factors on AMPK activity in the hypothalamus.

The recombinant adenoviruses expressed AMPK subunit messenger RNA primarily in ARH, VMH and DMH with very low-level expression in PVH and LH (Fig. 2a). Corresponding proteins were also detected in the ARH and VMH/DMH (Fig. 2b). Consistent with these results, immunohistochemistry with an antibody recognizing enhanced GFP (eGFP), also encoded in the $\alpha 2$ DN-AMPK adenovirus, showed many positive cells and fibres in VMH and DMH, the injection site of the adenoviruses (Fig. 2c). In ARH, several cells and many fibres were also positive. There were many positive fibres but no positive cells in PVH, LH and perifornical area.

In the first 2 days after adenovirus injection, all mice lost weight (Fig. 2d). However, mice expressing CA-AMPK regained weight more rapidly than controls. In contrast, mice with DN-AMPK lost more weight than controls or CA-AMPK-expressing mice and regained weight more slowly. These changes could at least partly be explained by alterations in food intake (Fig. 2e). Mice with DN-AMPK ate 0.3–0.5 g per day less than controls and those with CA-AMPK ate 0.3–0.5 g per day more starting at 4 days after adenovirus injections (Fig. 2e). Cumulative food intake over 8 days was also increased in CA-AMPK mice (50.5 ± 0.8 g, $n = 25$) compared to controls (47.8 ± 0.8 g, $n = 26$, $P < 0.05$) and was reduced in DN-AMPK mice (45.2 ± 0.6 g, $n = 27$, $P < 0.05$). Thus,

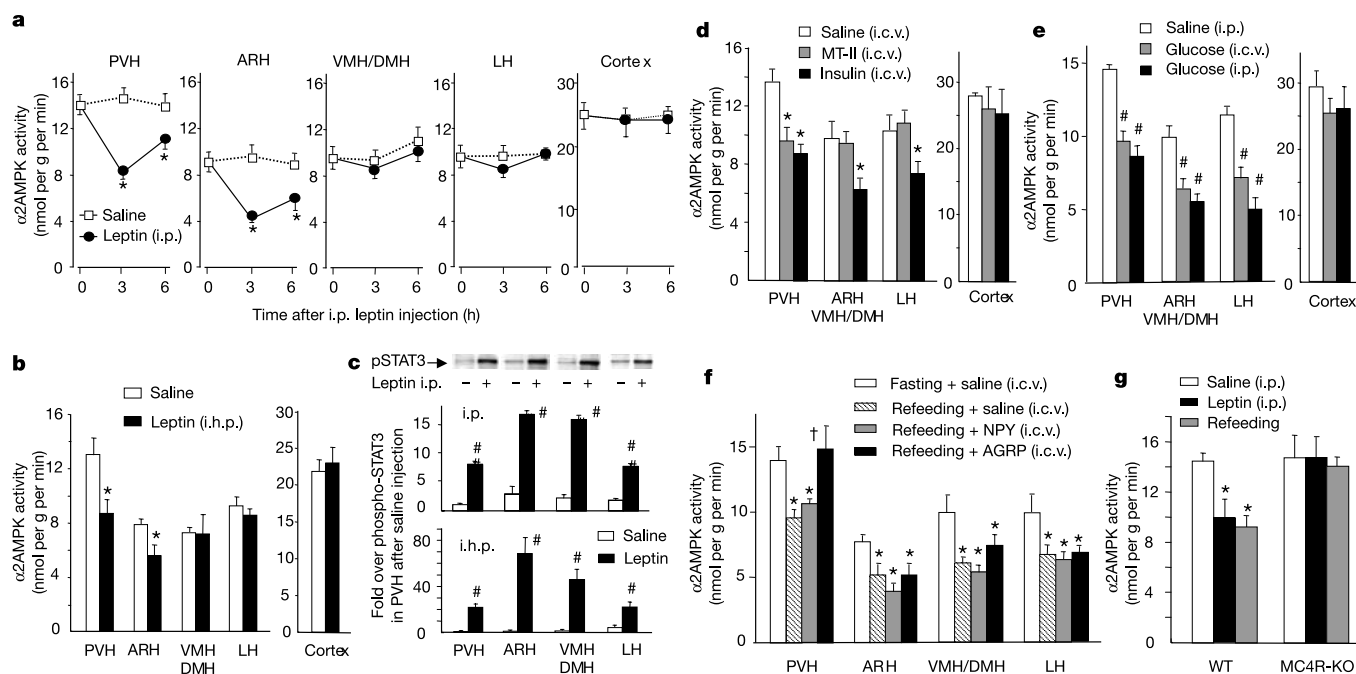


Figure 1 Leptin, insulin, MT-II and glucose decrease $\alpha 2$ AMPK activity in the hypothalamus. **a**, Leptin i.p. preferentially decreases $\alpha 2$ AMPK activity in PVH and ARH, but not in VMH/DMH, LH or cortex ($n = 10$ – 23 per group). **b**, $\alpha 2$ AMPK activity in the hypothalamus and cortex 3 h after intrahypothalamic injection (i.h.p.) of leptin ($n = 5$). **c**, Phosphorylation of STAT3 in the hypothalamus 3 h after i.p. or i.h.p. injection of leptin ($n = 5$). **d**, Injection (i.c.v.) of MT-II decreases $\alpha 2$ AMPK activity only in PVH 3 h after injection ($n = 8$). However, i.c.v. insulin decreases $\alpha 2$ AMPK activity in all hypothalamic regions 3 h after injection ($n = 8$). **e**, Injection (i.p. and i.c.v.) of glucose decreases $\alpha 2$ AMPK activity in all hypothalamic regions 1 h after injection ($n = 8$). * $P < 0.05$,

$P < 0.01$ versus saline. **f**, Refeeding (2 h) decreases $\alpha 2$ AMPK activity in all hypothalamic regions. AGRP (i.c.v.) but not NPY (i.c.v.) increases $\alpha 2$ AMPK activity in PVH to the fasting level 2 h after injection ($n = 5$ – 6). * $P < 0.05$ versus fasting plus saline. † $P < 0.05$ versus refeeding plus saline. **g**, Leptin (i.p.) and refeeding do not decrease $\alpha 2$ AMPK activity in PVH in MC4 receptor knockout (MC4R-KO, 6 weeks old) mice 3 h after injection or with refeeding ($n = 5$ – 6). Serum glucose in wild type (WT) and MC4R-KO was 5.3 ± 0.3 and 5.5 ± 0.3 mM, respectively, and serum insulin was 0.17 ± 0.03 and 0.26 ± 0.09 ng ml⁻¹, respectively. * $P < 0.05$ versus saline.

modulation of hypothalamic AMPK activity is sufficient to alter food intake and body weight.

To explore the mechanism(s) for these effects, we examined hypothalamic neuropeptide expression. Under *ad libitum* fed conditions, DN-AMPK decreased neuropeptide Y (NPY) and AGRP mRNA levels in ARH (Fig. 2f) whereas CA-AMPK had no effect. In contrast, CA-AMPK enhanced the fasting-induced increase in NPY (60% > fasting Null; $P < 0.01$) and AGRP mRNA levels (30% > fasting Null; $P < 0.05$). In mice expressing DN-AMPK, fasting increased NPY and AGRP expression similar to controls. Thus, DN- and CA-AMPK reciprocally regulate NPY and AGRP expression in ARH, depending on the feeding state, suggesting that high AMPK activity enhances orexigenic signals in the fasting state, whereas low AMPK activity suppresses these signals under *ad libitum* fed conditions. Consistent with these changes, CA-AMPK but not Null or DN-AMPK increased expression of melanin concentrating hormone (MCH), another orexigenic neuropeptide, in LH in response to fasting. This increase in MCH expression is probably secondary to changes in neuronal activity in the medial

hypothalamus including NPY/AGRP neurons, because we find very few eGFP-positive cells in LH (Fig. 2c). In contrast, pro-opiomelanocortin (POMC) expression in ARH responded normally to fasting in CA-AMPK-expressing mice. The lower POMC mRNA level in DN-AMPK-expressing mice under *ad libitum* conditions most probably reflects lowered body weight, food intake and serum leptin levels. These results indicate that POMC neurons respond physiologically in DN- and CA-AMPK-expressing mice, without augmentation or impairment of the fasting response.

To determine whether inhibition of AMPK activity in the hypothalamus is required for leptin's effect on food intake and body weight, we injected fasted mice expressing DN- or CA-AMPK in the medial hypothalamus with leptin. In control mice (adenovirus alone, Null), leptin decreased body weight by 2 g over 24 h compared to a slight increase in body weight in saline-injected control mice (Fig. 3a). Leptin also decreased body weight in mice expressing DN-AMPK. In contrast, in mice expressing CA-AMPK, body weight after saline injection was slightly greater than in saline-injected control (Null) mice, and leptin had a markedly attenuated

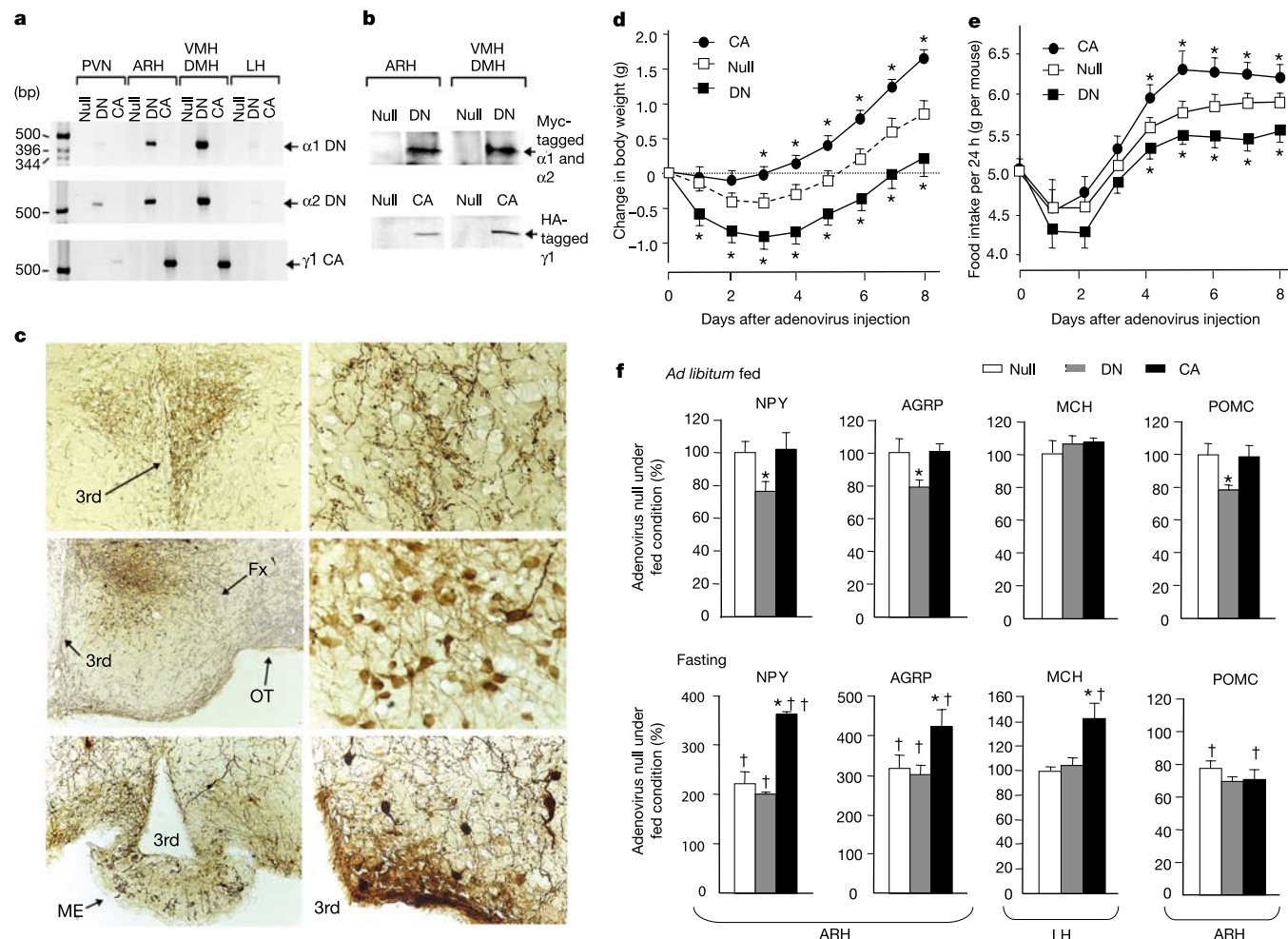


Figure 2 Modulation of AMPK activity in the hypothalamus is sufficient to alter body weight and food intake. **a**, Expression of DN-AMPK ($\alpha 1$ and $\alpha 2$) and CA-AMPK ($\gamma 1$) is primarily in the medial hypothalamus (ARH and VMH/DMH) with low expression in PVH and LH at 8 days after adenovirus injection. **b**, Protein expression of DN-AMPK ($\alpha 1$ and $\alpha 2$) and CA-AMPK ($\gamma 1$) in ARH and VMH/DMH detected with anti-Myc or anti-HA antibody at 8 days after injection. **c**, Immunohistochemistry with a specific antibody against eGFP. Upper-left: low magnification of PVH. Upper-right: high magnification of PVH. Middle-left: low magnification of ARH, VMH, DMH and LH. Middle-right: high magnification of VMH. Lower-left: intermediate magnification of ARH and ME (median eminence). Lower-right:

high magnification of ARH. 3rd, third ventricle; Fx, fornix; OT, optic tract. **d**, Changes in body weight after injection of DN-AMPK, CA-AMPK and adenovirus alone into the medial hypothalamus ($n = 25-27$). **e**, Daily food intake after injection of DN-AMPK, CA-AMPK or adenovirus alone into the medial hypothalamus ($n = 25-27$). **f**, DN- and CA-AMPK reciprocally regulate mRNA levels of NPY and AGRP in ARH depending on the feeding state ($N = 8-10$). * $P < 0.05$ versus adenovirus alone (Null). † $P < 0.05$, †† $P < 0.01$ versus the *ad libitum* fed condition for mice expressing the same adenovirus. Fasted and fed mRNA expression were measured in the same assay so that the levels can be directly compared. Null, adenovirus alone; DN, DN-AMPK; CA, CA-AMPK.

effect on body weight (Fig. 3a). Effects on food intake paralleled the effects on body weight. Leptin decreased food intake from 5.1 ± 0.2 g for 24 h at baseline to 2.9 ± 0.2 g in control mice (Fig. 3b). Food intake was decreased in mice with DN-AMPK at baseline and leptin further inhibited food intake to a level similar to leptin-treated control mice. However, in mice expressing CA-AMPK, food intake was increased at baseline and failed to be significantly suppressed by leptin. Thus, suppression of AMPK activity in the medial hypothalamus is necessary for leptin's anorectic and weight loss effects and lack of suppression causes leptin resistance.

In control mice (Null) in the fasted state, $\alpha 2$ AMPK activity in all hypothalamic regions was similar to the activity in fasted mice with no adenovirus treatment (Fig. 1a). Intrahypothalamic leptin in null mice decreased $\alpha 2$ AMPK activity in PVH and ARH (Fig. 3c), as was seen in mice with no adenovirus treatment (Fig. 1b). In mice with DN-AMPK, baseline $\alpha 2$ AMPK activity was decreased in PVH, ARH and VMH/DMH and leptin had no further effect. In fasted mice expressing CA-AMPK, baseline $\alpha 2$ AMPK was increased in VMH/DMH and leptin failed to decrease $\alpha 2$ AMPK activity in PVH, ARH or VMH/DMH (Fig. 3c). $\alpha 2$ AMPK protein level did not change (Supplementary Fig. 2).

We hypothesized that AMPK activity was not increased in the fasted state in some hypothalamic nuclei of mice expressing CA-AMPK because AMPK activity is already elevated with fasting. Hence we measured total ($\alpha 1$ and $\alpha 2$) AMPK activity after refeeding. In mice with adenovirus-null, refeeding decreased AMPK activity in the ARH and VMH/DMH to comparable levels as DN-AMPK (Fig. 3d). In contrast, in mice with CA-AMPK, total

AMPK activity in ARH and VMH/DMH was constitutively elevated in the refeed state. Consistent with the AMPK activity, food consumption during 3 h of refeeding was increased in mice with CA-AMPK and decreased in mice with DN-AMPK (Null: 1.54 ± 0.19 g per mouse, DN: 1.28 ± 0.10 , CA: 1.94 ± 0.16 , $P < 0.05$ for DN or CA versus Null).

The effects of DN- and CA-AMPK could not be explained by alterations in STAT3 tyrosine phosphorylation or protein level in the hypothalamus, because leptin's effect of increasing STAT3 phosphorylation was normal in ARH and VMH/DMH in mice with DN-AMPK and slightly enhanced in ARH in mice with CA-AMPK (Fig. 3e, Supplementary Fig. 2). In spite of this enhanced STAT3 phosphorylation in the ARH of CA-AMPK mice, leptin did not reduce food intake. This suggests that the early steps in leptin signalling are intact in the hypothalamus of mice expressing DN- or CA-AMPK, and that AMPK functions either downstream of, or in a parallel pathway to, STAT3 to modulate its effects on food intake.

Our data show that hypothalamic AMPK activity is suppressed by multiple anorexigenic factors and increased by the orexigenic peptide, AGRP. (In agreement with this, we note a recent report¹⁵, appearing online after we submitted the present Letter, showing that leptin inhibits, and ghrelin stimulates, hypothalamic AMPK activity.) We show that these effects are important in normal physiology because modulation of hypothalamic AMPK activity is sufficient to alter food intake and body weight (Fig. 2d, e). Furthermore, suppression of AMPK activity is necessary for the anorectic and weight loss effects of leptin (Fig. 3a, b). Thus, our data establish a novel pathway for the regulation of food intake which, in response to leptin, most probably acts in a coordinate

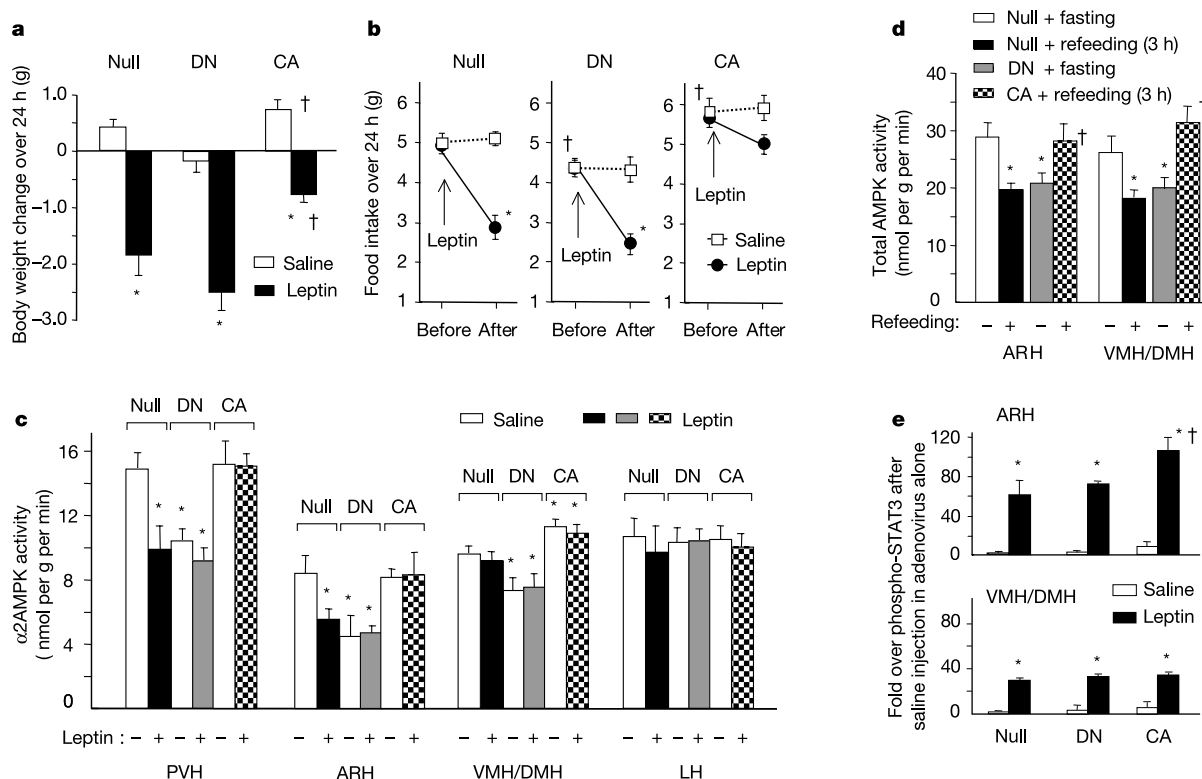


Figure 3 Effects of leptin are impaired with CA-AMPK in the hypothalamus. **a**, Body weight change 24 h after intrahypothalamic leptin ($n = 8-11$). * $P < 0.01$ versus saline injection. † $P < 0.05$ versus adenovirus alone (Null). **b**, Food intake for 24 h before and after intrahypothalamic leptin ($n = 8-11$). * $P < 0.01$ versus saline injection. † $P < 0.05$ versus food intake before leptin injection in Null. **c**, $\alpha 2$ AMPK activity in the hypothalamus in response to intrahypothalamic leptin (10 ng, 24 h and 3 h before mice were killed) at 8 days after the adenovirus injection ($n = 8-11$). Mice were fasted overnight. * $P < 0.05$

versus saline injection in Null. **d**, DN-AMPK decreases total AMPK activity in the fasting state (overnight) and CA-AMPK increases the activity in the refeed state (3 h) ($n = 5-6$). * $P < 0.05$ versus fasted Null. † $P < 0.05$ versus refeed Null. **e**, CA-AMPK does not interfere with phosphorylation of STAT3 in ARH and VMH/DMH in response to intrahypothalamic leptin ($n = 8-11$). * $P < 0.01$ versus saline injection. † $P < 0.05$ versus Null plus leptin injection.

manner with the STAT3¹⁶ and PI3 kinase^{17,18} pathways (Fig. 4).

All physiological anorexigenic signals that we examined inhibit $\alpha 2$ AMPK activity in the ARH and PVH. In the ARH, NPY and AGRP but not POMC neurons are affected (see model, Fig. 4). The possibility that another pathway in addition to STAT3 is critical for leptin's effects on expression of NPY and AGRP, but not POMC, was suggested in studies of a transgenic mouse expressing leptin receptors with a mutation in the tyrosyl residue required for STAT3 phosphorylation¹⁶. Similarly, inhibition of CPT1 in the hypothalamus in rats results in alterations in NPY and AGRP but not POMC expression¹⁹, also supporting the notion that these neuropeptides are regulated by different signalling pathways.

Our data suggest that alterations in AMPK activity in the PVH are mediated by the MC4 receptor, and are secondary to changes in activity of arcuate AGRP neurons (Fig. 4). AGRP competitively antagonizes the stimulatory effects of POMC neurons on the MC4 receptor¹². Furthermore, NPY functionally inhibits the MC4 receptor pathway in PVH²⁰. The effect of anorexigenic signals on AMPK activity in PVH appears to be mediated through MC4 receptors, as it can be mimicked with an MC3/4R agonist (Fig. 1d) and it is absent in MC4 receptor knockout mice (Fig. 1g). Although STAT3 and possibly PI3 kinase pathways may play a role in suppressing the

activity of NPY/AGRP neurons, we propose that decreased ARH AMPK activity enhances this suppression, leading to activation of MC4 receptor signalling in PVH neurons. MC4 receptor activation decreases AMPK activity in PVH, which probably further enhances neurotransmission required for regulation of food intake and energy balance. In addition, our data indicate that insulin, glucose and refeeding also decrease $\alpha 2$ AMPK activity in other hypothalamic regions, suggesting that these hormonal and nutritional signals recruit additional pathways that may regulate energy balance and have other physiological functions.

The effect of AMPK may be through transcriptional effects¹, actions on ion channels^{21,22} or changes in cytosolic Ca^{2+} (ref. 23). Leptin, MT-II, insulin and glucose each alter the activity of K_{ATP} channels^{24,25} and other ion channels in neurons²⁰, some of which regulate cytosolic Ca^{2+} concentrations. These results are also consistent with increased hypothalamic malonyl-CoA levels^{26,27} and direct inhibition of hypothalamic CPT1¹⁹ decreasing food intake. Thus, the evolutionarily conserved energy-sensing enzyme, AMPK, regulates energy balance in higher organisms not only by altering metabolism^{1,28,29}, but also by responding to nutritional and hormonal signals governing food intake. Understanding the integration of this pathway with other afferent signals could lead to new approaches to prevent or reverse obesity. □

Methods

Animals

Male FVB mice (aged 7–9 weeks) were housed in a temperature-controlled environment with a 14/10-h light/dark cycle. Mice were given chow (Formulab 5008; Farmer's Exchange) and water freely before beginning experiments. Some mice were implanted with chronic cannulas in the medial hypothalamus bilaterally 2 weeks before experiments. Other mice were implanted with an unilateral chronic cannula in the lateral ventricle. Leptin (10 ng) was injected through the cannula in the medial hypothalamus. Insulin (1 pmol), MT-II (30 pmol), NPY (10 nmol) and AGRP (10 nmol) were injected i.c.v. through the cannula in the lateral ventricle. In some mice, including MC4 receptor knockout mice, leptin (3 mg kg⁻¹) or saline was injected i.p. In other mice, adenoviruses expressing dominant negative $\alpha 1^{13}$ or $\alpha 2$ AMPK (Supplementary Methods), constitutively active $\gamma 1$ (Supplementary Methods), a control adenoviral vector without DNA insert were injected into the medial hypothalamus bilaterally 2 weeks after catheter implantation. Body weight and food intake were monitored daily. Mice were killed by decapitation and hypothalamic nuclei and parietal cortex were quickly dissected. All assays were performed on hypothalamic regions from individual mice. Stereotaxic coordinates for intrahypothalamic cannula and injections are described in Supplementary Methods.

Dissection of hypothalamic regions and cortex

Each hypothalamic region and parietal cortex were dissected from 1-mm-thick sagittal sections of fresh brain. PVH, ARH, VMH plus DMH, and parietal cortex were dissected from the first sections from the midline of the brain, and LH was dissected from the next lateral sections. In some experiments, ARH and VMH plus DMH were dissected together. Coordinates for each hypothalamic region are described in Supplementary Methods.

Measurement of AMPK activity

To measure $\alpha 1$ and $\alpha 2$ isoform-specific AMPK activity in the hypothalamic regions and cortex, we immunoprecipitated each AMPK from lysates from individual mice (40–50 μ g of protein) with specific antibodies against the $\alpha 1$ - (ref. 30) and $\alpha 2$ -subunits (ref. 30) bound to protein-G sepharose beads. To measure total AMPK activity, we used a specific antibody¹³ against the β -subunits bound to protein-A and -G sepharose beads. The kinase activity of the immunoprecipitates was measured using 'SAMS' peptide and [γ -³²P]ATP³¹.

Measurement of mRNA levels of neuropeptides

mRNA levels of TRH, AGRP, NPY, POMC and MCH were quantified with real time PCR, as described in Supplementary Methods. All data were expressed as a ratio of the neuropeptide mRNA to cyclophilin mRNA. The data are shown as a percentage of the ratio in null fed mice.

Immunohistochemistry with eGFP antibody

Mice were injected with an adenovirus expressing both DN- $\alpha 2$ AMPK and eGFP (2.5 $\times 10^6$ p.f.u. per 0.1 μ l, Supplementary Methods) into the medial hypothalamus bilaterally, using brain cannula implanted with stereotaxic coordinates as described above. Six days after adenovirus injection, animals underwent transcardiac formalin fixation and immunohistochemistry of brain was performed with an eGFP antibody (1:20,000, Molecular Probes).

Western blot analysis

Phosphorylation of STAT3 in hypothalamic regions and cortex was determined with 10% SDS acrylamide gels using an antibody against the phospho-tyrosine705 of STAT3 (Cell Signalling). The protein levels of $\alpha 1$ and $\alpha 2$ AMPK³⁰ and STAT3 (Santa Cruz)

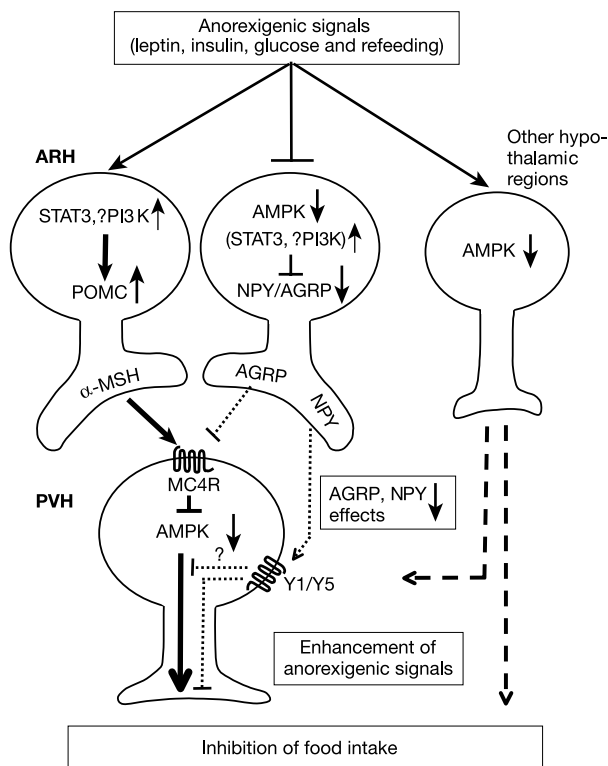


Figure 4 Proposed model for role of AMPK in anorexigenic signalling in the hypothalamus. Anorexigenic signals activate POMC neurons in ARH (arcuate hypothalamus) via STAT3 and possibly also PI3 kinase, generating a second anorexigenic signal mediated by α -melanocyte stimulating hormone (α -MSH). In contrast, the anorexigenic signals suppress the activity of NPY/AGRP neurons, partly via STAT3 and possibly also PI3 kinase, and decrease AMPK activity in these neurons. Decreased AMPK activity enhances the suppression of NPY/AGRP effects, leading to activation of MC4 receptor signalling in PVH (paraventricular hypothalamus) neurons. MC4 receptor activation decreases AMPK activity in PVH, which probably further enhances neurotransmission required for regulation of food intake and energy balance. Decreased NPY signalling in PVH functionally enhances the MC4 receptor signalling pathway. In addition, decreased AMPK activity in other hypothalamic regions may enhance the MC4 receptor signalling pathway by projections to the ARH or PVH (for example, NPY neurons in DMH) and recruit additional pathways that may regulate food intake.

in the hypothalamic regions were determined.

To detect protein expression of $\alpha 1$ and $\alpha 2$ DN-AMPK and $\gamma 1$ CA-AMPK in the hypothalamus after injection of adenoviruses, we immunoprecipitated AMPK from hypothalamic lysates (500 μ g of protein pooled from 5–6 animals) with a polyclonal antiserum recognizing the $\alpha 1$, $\alpha 2$, $\beta 1$, $\beta 2$ and $\gamma 1$ subunits of AMPK (for CA-AMPK) (gift from D. Carling) or this antiserum combined with sheep $\alpha 1$ and $\alpha 2$ antiserum (for DN-AMPK) bound to protein-A and -G sepharose beads, and blotted with monoclonal antibodies against the c-Myc tag (for $\alpha 1$ and $\alpha 2$ DN-AMPK) (9B11, Cell Signalling) or the HA tag (for $\gamma 1$ CA-AMPK) (Roche).

Detection of mRNA of DN- and CA-AMPK

Total RNA was isolated from PVH, ARH, VMH/DMH and LH by TriReagent (Molecular Research Center). First-strand cDNA was synthesized from 2 μ g of total RNA using reverse transcriptase (Ambion) primed by random decamer. PCR amplification of Myc-tagged $\alpha 1$ and $\alpha 2$ DN-AMPK and HA-tagged $\gamma 1$ CA-AMPK was performed with Platinum Tag DNA polymerase (Invitrogen). The conditions of PCR and design of the primers are described in Supplementary Methods.

Statistical analysis

All values are mean $\pm$ s.e.m. Data were evaluated by factorial analysis of variance and the Newman-Keuls multiple range test.

Received 22 December 2003; accepted 27 February 2004; doi:10.1038/nature02440.

Published online 17 March 2004.

- Hardie, D. G., Scott, J. W., Pan, D. A. & Hudson, E. R. Management of cellular energy by the AMP-activated protein kinase system. *FEBS Lett.* **546**, 113–120 (2003).
- Schwartz, M. W. *et al.* Central nervous system control of food intake. *Nature* **404**, 661–671 (2000).
- Friedman, J. M. & Halaas, J. L. Leptin and the regulation of body weight in mammals. *Nature* **395**, 763–770 (1998).
- Brüning, J. C. *et al.* Role of brain insulin receptor in control of body weight and reproduction. *Science* **289**, 2122–2125 (2000).
- Levin, B. E. Glucosensing neurons do more than just sense glucose. *Int. J. Obes. Relat. Metab. Disord. Suppl.* **5**, S68–S72 (2001).
- Obici, S. *et al.* Central administration of oleic acid inhibits glucose production and food intake. *Diabetes* **5**, 271–275 (2002).
- Hawley, S. A. *et al.* Complexes between the LKB1 tumor suppressor, STRADA/ β and MO25/ α / β are upstream kinases in the AMP-activated protein kinase cascade. *J. Biol.* **2**(28), 1–16 (2003).
- Turnley, A. M. *et al.* Cellular distribution and developmental expression of AMP-activated protein kinase isoforms in mouse central nervous system. *J. Neurochem.* **72**, 1707–1716 (1999).
- Culmsee, C., Monnig, J., Kemp, B. E. & Mattson, M. P. AMP-activated protein kinase is highly expressed in neurons in the developing rat brain and promotes neuronal survival following glucose deprivation. *J. Mol. Neurosci.* **17**, 45–58 (2001).
- Elmqvist, J. K., Elias, C. F. & Saper, C. B. From lesions to leptin: hypothalamic control of food intake and body weight. *Neuron* **22**, 221–232 (1999).
- Obici, S., Zhang, B. B., Karkanas, G. & Rossetti, L. Hypothalamic insulin signaling is required for inhibition of glucose production. *Nature Med.* **8**, 1376–1382 (2002).
- Ollmann, M. M. *et al.* Antagonism of central melanocortin receptors *in vitro* and *in vivo* by agouti-related protein. *Science* **278**, 135–138 (1997).
- Woods, A. *et al.* Characterization of the role of AMP-activated protein kinase in the regulation of glucose-activated gene expression using constitutively active and dominant negative forms of the kinase. *Mol. Cell. Biol.* **20**, 6704–6711 (2000).
- Viollet, B. *et al.* The AMP-activated protein kinase $\alpha 2$ catalytic subunit controls whole-body insulin sensitivity. *J. Clin. Invest.* **111**, 91–98 (2003).
- Andersson, U. *et al.* AMP-activated protein kinase plays a role in the control of food intake. *J. Biol. Chem.* published online 23 January 2004 (doi:10.1074/jbc.C300557200).
- Bates, S. H. *et al.* STAT3 signalling is required for leptin regulation of energy balance but not reproduction. *Nature* **421**, 856–859 (2003).
- Niswender, K. D. *et al.* Intracellular signalling. Key enzyme in leptin-induced anorexia. *Nature* **413**, 794–795 (2001).
- Zhao, A.-Z. *et al.* A phosphatidylinositol 3-kinase-phosphodiesterase 3B-cyclic AMP in hypothalamic action of leptin on feeding. *Nature Neurosci.* **5**, 727–728 (2002).
- Obici, S. *et al.* Inhibition of hypothalamic carnitine palmitoyltransferase-1 decreases food intake and glucose production. *Nature Med.* **9**, 756–761 (2003).
- Cowley, M. A. *et al.* Integration of NPY, AGRP, and melanocortin signals in the hypothalamic paraventricular nucleus: evidence of a cellular basis for the adipostat. *Neuron* **24**, 155–163 (1999).
- Light, P. E., Wallace, C. H. R. & Dyck, J. R. B. Constitutively active adenosine monophosphate-activated protein kinase regulates voltage-gated sodium channels. *Circulation* **107**, 1962–1965 (2003).
- Hallows, K. R. *et al.* Inhibition of cystic fibrosis transmembrane conductance regulator by novel interaction with the metabolic sensor AMP-activated protein kinase. *J. Clin. Invest.* **12**, 1711–1721 (2000).
- da Silva Xavier, G. *et al.* Role for AMP-activated protein kinase in glucose-stimulated insulin secretion and preproinsulin gene expression. *Biochem. J.* **371**, 761–774 (2003).
- Spaniswick, D. *et al.* Leptin inhibits hypothalamic neurons by activation of ATP-sensitive potassium channels. *Nature* **390**, 521–525 (1997).
- Spaniswick, D. *et al.* Insulin activates ATP-sensitive K^+ channels in hypothalamic neurons of lean, but not obese rats. *Nature Neurosci.* **3**, 757–758 (2000).
- Loftus, T. M. *et al.* Reduced food intake and body weight in mice treated with fatty acid synthase inhibitors. *Science* **288**, 2379–2381 (2000).
- Hu, Z., Cha, S. H., Chohan, S. & Lane, D. Hypothalamic malonyl-CoA as a mediator of feeding behavior. *Proc. Natl Acad. Sci. USA* **100**, 12624–12629 (2003).
- Ruderman, N. B., Saha, A. K., Vavvas, E. & Witters, L. A. Malonyl-CoA, fuel sensing, and insulin resistance. *Am. J. Physiol.* **276**, E1–E18 (1999).
- Minokoshi, Y. *et al.* Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* **415**, 339–343 (2002).

- Woods, S. *et al.* The $\alpha 1$ and $\alpha 2$ isoforms of the AMP-activated protein kinase have similar activities in rat liver but exhibit differences in substrate specificity *in vitro*. *FEBS Lett.* **397**, 347–351 (1996).
- Hayashi, T. *et al.* Metabolic stress and altered glucose transport. Activation of AMP-activated protein kinase as a unifying coupling mechanism. *Diabetes* **49**, 527–531 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank D. Carling for reagents and advice, J. K. Elmquist and B. B. Lowell for discussions and providing MC4R-KO mice, and C. J. Aschkenasi, C.-Y. Zhang, O. Boss, J. Yu and N. Balthasar for MC4R-KO mice. This work was supported by NIH grants (B.B.K. and M.J.B.), an EASD-ADA and Bettencourt-Schueller Foundation Fellowship (T.A.), AMPDIAMET (P.F.) and the American Diabetes Association (B.B.K. and Y.B.K.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.B.K. (bkahn@bidmc.harvard.edu).

From molecular noise to behavioural variability in a single bacterium

Ekaterina Korobkova^{1*}, Thierry Emonet^{1*}, Jose M. G. Vilar^{2†}, Thomas S. Shimizu^{3†} & Philippe Cluzel¹

¹The Institute for Biophysical Dynamics and the James Franck Institute, The University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA
²The Rockefeller University, Box 34, 1230 York Avenue, New York, New York 10021, USA

³Laboratory for Bioinformatics, Institute for Advanced Biosciences, Keio University Fujisawa, 252-8520, Japan

* These authors contributed equally to this work

† Present addresses: Computational Biology Center, Memorial Sloan-Kettering Cancer Center, 307 East 63rd Street, New York, New York 10021, USA (J.M.G.V.); Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA (T.S.S.)

The chemotaxis network that governs the motion of *Escherichia coli* has long been studied to gain a general understanding of signal transduction. Although this pathway is composed of just a few components, it exhibits some essential characteristics of biological complexity, such as adaptation and response to environmental signals¹. In studying intracellular networks, most experiments and mathematical models^{2–5} have assumed that network properties can be inferred from population measurements. However, this approach masks underlying temporal fluctuations of intracellular signalling events. We have inferred fundamental properties of the chemotaxis network from a noise analysis of behavioural variations in individual bacteria. Here we show that certain properties established by population measurements, such as adapted states, are not conserved at the single-cell level: for timescales ranging from seconds to several minutes, the behaviour of non-stimulated cells exhibit temporal variations much larger than the expected statistical fluctuations. We find that the signalling network itself causes this noise and identify the molecular events that produce it. Small changes in the concentration of one key network component suppress temporal behavioural variability, suggesting that such variability is a selected property of this adaptive system.

At the level of populations, it is well-established that the chemotaxis network produces a steady output in the absence of external stimuli, generally referred to as 'adapted steady-states'. This mechanism³ allows bacteria to maintain their steady-state behaviour independently of the absolute concentration of chemo-effectors in their environment⁵. Because the network's output from individual cells is noisy (that is, it fluctuates randomly), it is standard practice

to average responses across a population of cells, which eliminates part of the information required to understand how an individual cell performs basic computations^{6,7}. But noise is not always a nuisance: it can carry important information about intracellular signalling events^{8–10}. We investigated how the behaviour of an individual bacterium of *E. coli* in a homogeneous environment fluctuates with time. In particular, we asked whether there are specific molecular events that could cause temporal behavioural variability in an individual cell.

We monitored the switching events of individual flagellar motors¹¹ from non-stimulated cells in a medium in which attractant was not present. Bacteria were immobilized onto microscope slides and flagella were marked with micro-beads to visualize their rotation with dark-field illumination¹². Binary time series constructed from the clockwise (CW) and the counterclockwise (CCW) rotations of a single motor defined the chemotaxis network output¹¹ (Fig. 1a). The CW bias is the fraction of time that a motor spends rotating in the CW direction^{11,13} (Fig. 1b). We studied the temporal variations of the CW bias by carrying out spectral analysis of the binary time series generated by switching events of individual bacterial motors. In this approach, the power spectrum is a measure of the amplitude of the fluctuations of the CW bias at a given timescale. The standard assumption has been that switching events are independent and governed by a Poisson process, which implies that the CW and CCW time intervals are uncorrelated and exponentially distributed¹³. Under these assumptions, the power spectrum would exhibit a flat profile over timescales larger than the typical switching time (Supplementary Information).

We recorded a 170-min-long time series of switching events from an individual motor of a RP437 wild-type bacterium, which was immersed in a medium in which no attractant was present and that did not support growth. Unexpectedly, the corresponding power spectrum exhibited a growing profile up to 15 min ($\sim 10^3$ s, Fig. 2a). The fact that the power spectrum is not flat indicates that the CCW and CW intervals are either not 'independently and identically distributed (IID)', or IID but not exponentially distributed. It also means that the variability of the CW bias over these timescales is larger than that expected from a motor with exponentially distributed intervals and similar mean switching frequency. Because not all the cells exhibited the same temporal variability, the slopes of power spectra varied from cell to cell, but all fell into the open interval (0, 1). Consequently, we computed 222 power spectra from clonal individual RP437 wild-type cells and summarized their average trend in Fig. 2b. We found that the trend of 222 individual

power spectra exhibited a growing profile as frequencies decreased. The temporal variability of the network output, the CW bias, was larger than that expected from uncorrelated and exponentially distributed CW and CCW intervals^{14,15}.

To reconcile our data with ensemble measurements, we averaged together the previous 222 binary time series from individual cells before spectral analysis so that we could study the nature of the fluctuations of the bias of a population. The resulting power spectrum exhibited a flat profile at any timescale greater than seconds, indicating that the temporal variations of the bias obtained from a population could be produced by uncorrelated, exponentially distributed CW and CCW intervals¹³ (Fig. 2b, inset). In this system, the time average of single-cell behaviour differs from population behaviour for timescales ranging from few to 10^3 seconds.

The chemotaxis network is a phosphoryl cascade that controls the concentration of the phosphorylated form of a signalling molecule, the soluble response regulator CheY¹ (Supplementary Information). The phosphorylated form, CheY-P, binds preferentially to the cytoplasmic base of the motors. When the concentration of CheY-P increases, motors spend more time spinning clockwise. To investigate the molecular origin of behavioural variability of single PS2001 mutant cells, we used the activated CheYD13K mutant, which mimics the effect of CheY-P but does not need to be phosphorylated to be active⁴. The active CheYD13K signalling molecule was expressed so that the mean CW bias from the population would be at about the wild-type level. The power spectrum associated with the CheYD13K mutant provided the spectral characteristics of the bacterial motor (Fig. 2b). At short timescales the spectrum was similar to wild type and peaked at about one second. For longer timescales, the data showed that when the concentration of the active form of the signalling molecule was not regulated by the chemotaxis network, but stably expressed from an inducible plasmid, the noise level was much lower than in wild-type cells (Fig. 2b). This result suggests that temporal behavioural variability in a wild-type cell emerges from the signalling processes taking place in the network itself.

We characterized the underlying statistical nature of the distributions of the CW and CCW switching events in the PS2001 mutant and wild-type cells. An individual wild-type cell exhibits a distribution of short CW length intervals that is dominated by an exponential behaviour (Fig. 2c). However, we found that the distribution of CCW intervals from several wild-type cells deviates significantly from exponential behaviour^{14,16} (Fig. 2c). In particular, this distribution could instead be approximated at long timescales

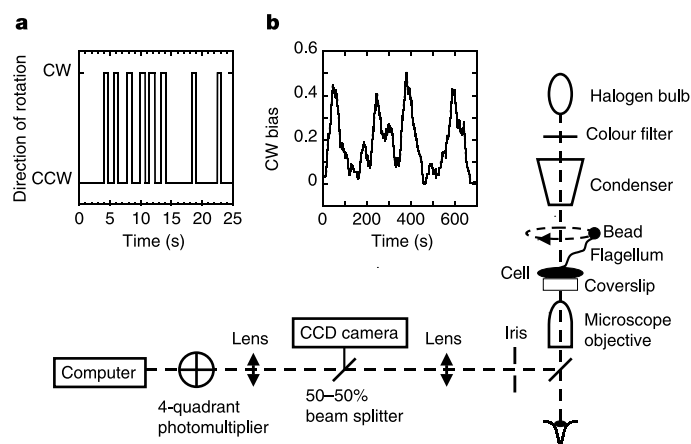


Figure 1 Schematic view of the apparatus. Cells were specifically attached onto a microscope slide. Latex beads (0.5 μ m, Polysciences) were used as markers to visualize single rotating flagella¹² on an inverted microscope with dark-field illumination. A four-quadrant photomultiplier (PMT) was used to record the trajectory of the bead. Typical

assays lasted from 10 to 20 min. **a**, Binary time series, indicating the direction of rotation. **b**, CW bias versus time. The bias was computed as the fraction of time the bead span clockwise within a 30 s moving window.

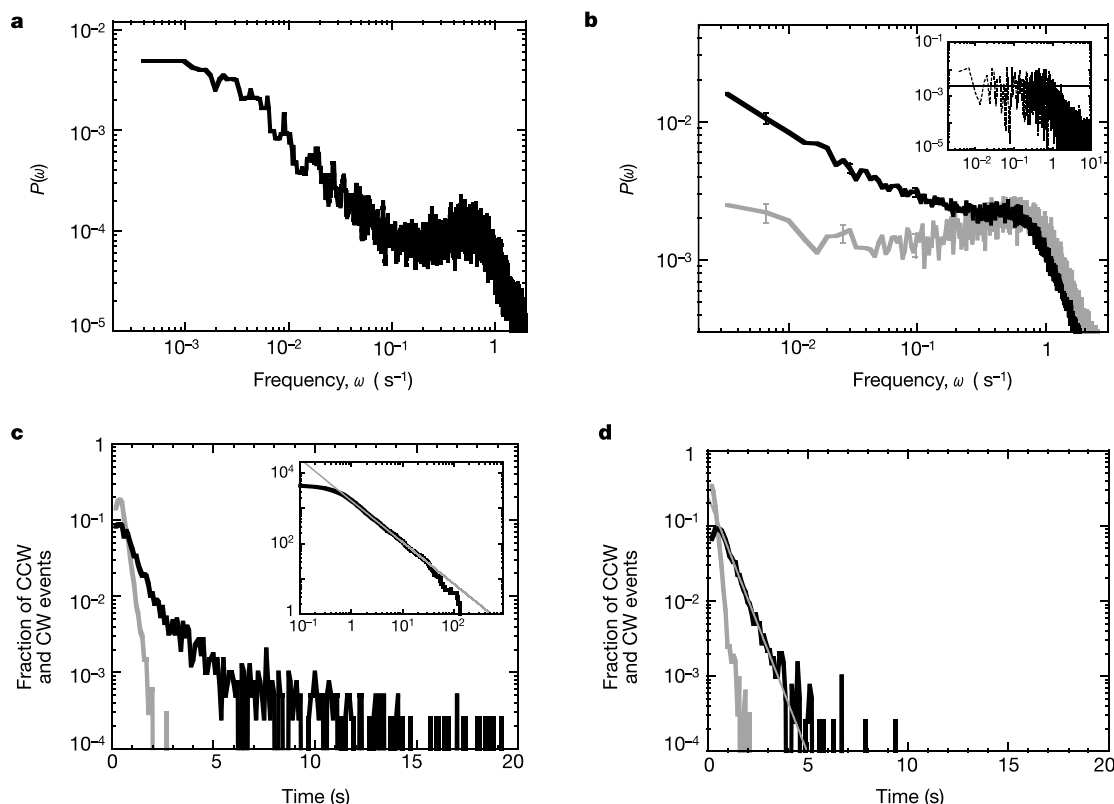


Figure 2 Noise of the chemotaxis network. **a**, Power spectrum of the network output from one non-stimulated individual wild-type cell. **b**, Mean spectrum computed by averaging the spectra from wild-type and mutant cells. Black line, mean spectrum from 40 wild-type cells; grey line, spectral characteristics of the bacterial motor. Mean power spectrum of individual power spectra from 16 PS2001 mutant cells complemented by the inducible *cheYD13K* gene ([IPTG] = 25, 30 μ M). The error bars (grey and black line) show the

standard error. Inset, power spectrum of the CW bias of a population of 40 cells.

c, Distribution of CW (grey) and CCW (black) intervals from the cell in **a**. Inset, cumulative distribution of the same CCW intervals (black line). Power law with an exponent of approximately -1.2 (grey straight line). **d**, CW and CCW interval distributions from eight mutants (PS2001) expressing *CheYD13K* (CW bias ranging from 0.2 to 0.3).

by a power law (Fig. 2c, inset). For mutant cells in which the motor output was decoupled from the activity of the signalling pathway, both CW and CCW intervals were exponentially distributed (Fig. 2d).

To determine the specific molecular events that account for the observed temporal variability of the CW bias, we focused on receptor methylation (Supplementary Information), which has been shown through population measurements to be a key determinant of the steady-state motor bias^{3,5,17}. We investigated how the temporal behavioural variability noted above depends on the concentration of the methyltransferase CheR, which adds methyl groups at multiple receptor residues. The *cheR*-deleted RP4968 mutant was complemented with a *lac*-inducible low-copy plasmid expressing the CheR protein⁵. Intracellular concentration, [CheR], was varied over an approximately tenfold range by induction with increasing amounts of isopropyl- β -D-thiogalactoside, [IPTG] (Methods). At low [CheR] ([IPTG] = 0 μ M), wild-type levels of behavioural variability were recovered. At timescales greater than seconds, the power spectrum of the network output of individual cells exhibited a growing profile similar to that of wild-type cells (Fig. 3a). Surprisingly, as [CheR] was increased to about two times the wild-type concentration ([IPTG] = 1 μ M), the power spectrum exhibited a weaker slope, which faded away for values of [CheR] $\geq$ four times the wild-type level (Fig. 3a). Similarly, as [CheR] was increased, the initially power-law-distributed CCW intervals segued towards exponential behaviour (Fig. 3b). Therefore, the temporal behavioural variability could be reduced and furthermore suppressed by increasing [CheR]. For individual wild-type cells with low [CheR], the output bias exhibited temporal variations

greater than statistical fluctuations expected from an exponential distribution of uncorrelated CW and CCW intervals, whereas for slightly higher [CheR] this noise was markedly suppressed (Fig. 3c).

To gain further insights into the signalling pathway, we simulated the chemical reactions between the cytosolic chemotaxis proteins (CheR, CheB, CheY and CheZ) and the receptor-kinase complexes (Supplementary Information). To capture time correlations in a reacting system, we performed stochastic simulations using the StochSim package⁶ (Methods). We simulated the temporal fluctuations in CheY-P concentration and computed the corresponding power spectra for various values of [CheR]. For approximately wild-type levels of [CheR], the wild-type-like spectrum was qualitatively recovered (Fig. 4a). The fluctuations of concentration of the active signalling molecules (CheY-P) exhibit a strong correlated noise. Moreover, the amplitude of these fluctuations could be modulated by the methylation and demethylation of the receptors catalysed by CheR and CheB (Supplementary Information). Because CheR works at saturation, the number of active receptors increases nonlinearly in an ultra-sensitive^{19,20} fashion with [CheR] (data not shown). At wild-type concentrations of CheR, about half of the receptors are active and the associated fluctuations of [CheY-P] are maximal^{21,22}. For higher [CheR], four times the wild-type level or more, almost all the receptors are active and the associated fluctuations of [CheY-P] are smaller (Fig. 4b).

To support the hypothesis that the methylation process controls temporal behavioural variability, we analysed the noise from the RP8610 (RP8610 strain and pRR27 plasmid; J. S. Parkinson, personal communication) mutant deleted for the *cheR*, *cheB* and all

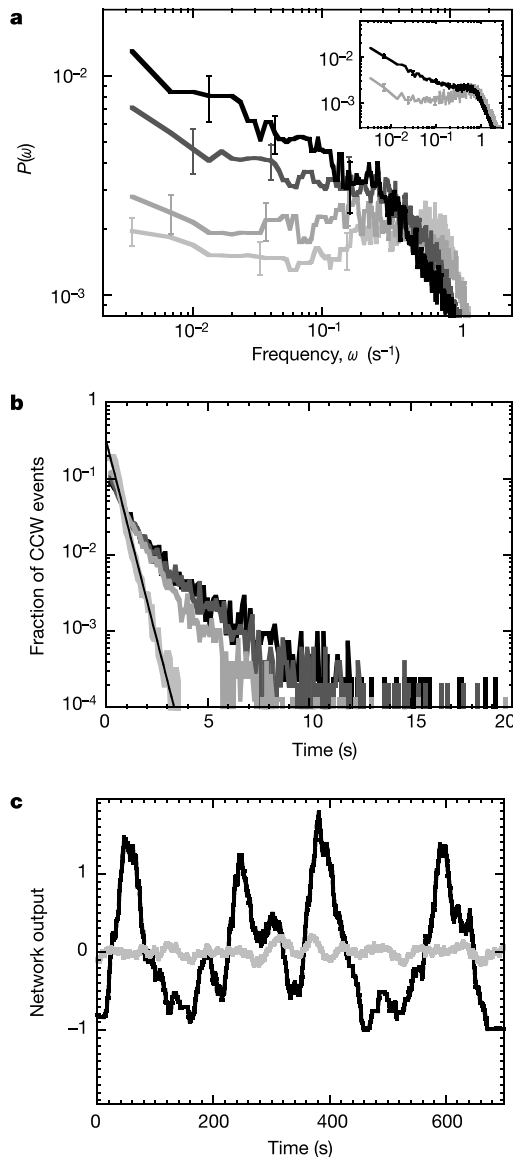


Figure 3 Behavioural variability as a function of [CheR]. **a**, Average power spectra of the motors, switching events from RP4968 cells versus [CheR] levels. Black, wild-type level of CheR ([IPTG] = 0 μ M); dark grey, twofold wild-type level ([IPTG] = 1 μ M); grey, fourfold wild-type level ([IPTG] = 5 μ M); light grey, tenfold wild-type level ([IPTG] = 30 μ M). Error bars show the standard error. Inset, effect of fixed methylation level on behavioural variability; black, as in Fig. 2a; grey, spectra from RP 8610 mutant cells complemented with Tsr mutant serine receptors. **b**, CCW interval distributions versus [CheR] (same cells and conditions as in panel a). Exponential fit (straight black line). **c**, Network output signals from a wild-type and a mutant cell. Black, wild-type cell with a mean CW bias of 0.2 and a switching frequency of 0.4 s⁻¹ (same cell as in Fig. 2a); grey, RP4968 mutant (concentration ten times the wild-type level) with a mean CW bias of 0.4 and a switching frequency of 0.9 s⁻¹ (same cell as a; light grey). The network output is defined as $[\text{bias} - <\text{bias} >] / <\text{bias} >$.

chemoreceptor genes. In this strain, mutant serine receptors that mimic the activity of receptors with a fixed level of methylation were stably expressed from a *lac*-inducible vector (Methods). The resulting variability was found to be smaller than in wild-type cells and the associated power spectra were identical to the spectral characteristics of the motor when uncoupled from the network (Fig. 3a, inset; compare with Fig. 2b). This result suggests that the slow methylation process of the receptors contributes to the observed behavioural variability.

It is conceivable that such variability, and in particular the power-

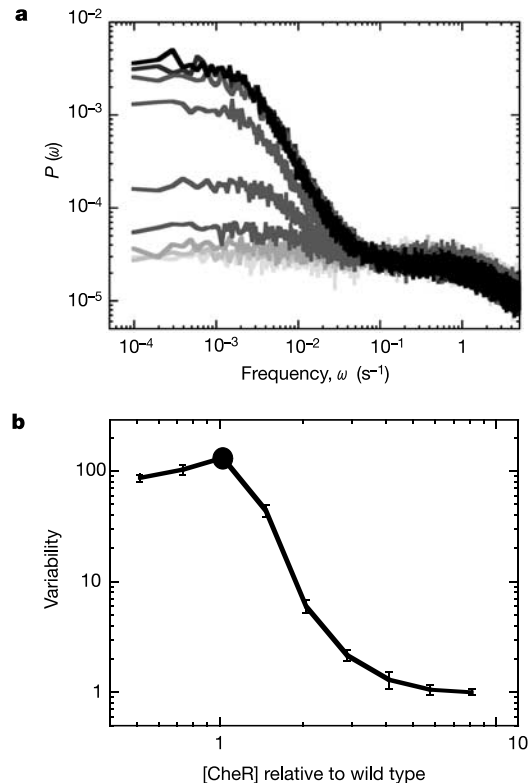


Figure 4 Simulated variability of the chemotaxis network. **a**, Power spectra of simulated temporal variations of the signalling molecule concentration [CheYp] for increasing [CheR]. [CheR] was increased from 0.5 (black) to eight times the wild-type level (light grey). In the frequency interval 0.1–1 Hz, the spectra displayed a flat region robust to changes in [CheR]. **b**, Variability of the network output [CheYp] versus [CheR]. Wild-type cell as in ref. 6 (filled circle). The variability was defined as the power of the output signal (for a chosen [CheR]) normalized by the power of the network output for a [CheR] eight times the wild-type level. Means of the power and standard deviations were computed for a frequency interval (10⁻³, 10⁻⁴ Hz).

law behaviour of CCW intervals, will be reflected in the run-length distribution of individual swimming cells. A power-law distribution of the run lengths may provide bacteria with an optimal search strategy to adapt to a complex environment with a sparse distribution of nutrients²³ (Supplementary Information).

Three decades ago, Spudich and Koshland established the existence of ‘non-genetic individuality’ in clonal populations of *Salmonella typhimurium*²⁴. We analysed the molecular origin of the temporal variations in signalling within individual bacteria. Both experimental and simulation data showed that within individual bacteria, molecular noise emerges as a tunable source of behavioural variability (Figs 3c and 4b). If the relative concentration of a key chemotaxis protein ([CheR]) had been slightly higher, any wild-type cell would have exhibited a steadier behaviour. Our results revealed that such a regime was not selected for in the chemotaxis system. On the contrary, the presence of a large temporal behavioural variability in this simple sensory system^{19,21,22} appears to be the manifestation of the ‘adapted’ state of wild-type cells (Fig. 4b). Considering the ubiquity in nature of signalling pathways with similar design principles¹⁹, it would be surprising not to find that molecular noise in other signal transduction networks is also a selectable source of cell fate variability. □

Methods

Bacterial strains and plasmids

Cells were grown from an overnight culture in tryptone broth at 30 °C and then harvested (optical density = 0.5 at 595 nm). Cells were then washed and suspended in minimal

medium (7.6 mM (NH₄)₂SO₄, 2 mM MgSO₄, 20 μM FeSO₄, 0.1 mM EDTA, 0.1 mM L-methionine, 60 mM potassium phosphate pH 6.8). PS2001 and DeltaR RP4968 mutants were grown with various IPTG concentrations. CheYD13K was expressed from pMS164 (ref. 5). The average [CheR] (expressed from pUA4) was estimated using the relationship between [IPTG] and [CheR] assessed by immunoblots in ref. 5.

PS2001 strain

This strain is deleted for *CheB*, *CheZ* and *CheY*. The strain was transformed with an inducible *lac* promoter pMS164 (ref. 4) plasmid expressing a *cheYD13K* gene. The CheYD13K mutant acts like CheY-P but does not need to be phosphorylated to be active. The active CheYD13K signalling molecule was expressed from the *lac*-inducible vector pMS164 (with [IPTG] = 25–30 μM).

RP4968 strain

cheR is deleted in this strain. Cells were complemented with a *lac*-inducible low-copy plasmid pUA4 (ref. 5) expressing the CheR protein. CW bias was controlled by the induction level of CheR.

RP8610 strain

This strain carries a complete deletion of *cheR*, *cheB*, *tsr*, *tar*, *tap* and *trg*. The strain was transformed with an inducible *lac* promoter pRR27 plasmid expressing the *Tsr* mutant gene²⁰. The methylation sites of this *Tsr* mutant are QQQE. Varying the *Tsr* expression allowed us to set the steady-state bias at any value. In Fig. 3a (inset), the bias was adjusted to about wild-type level.

Simulations

The reaction system was based on a two-state model¹⁸ of receptor activation, and its parameter values used to simulate the wild-type behaviour were identical to those described in ref. 6. In Fig. 4, [CheR] was incremented from its wild-type value in factors of $\sqrt{2}$. For each value of [CheR], 63 independent time sequences of CheYp (each 170 min long) were generated and normalized by their standard deviation. The average of the 63 corresponding power spectra was plotted for each [CheR].

Received 30 August 2003; accepted 5 February 2004; doi:10.1038/nature02404.

- Bourret, R. B. & Stock, A. M. Molecular information processing: lessons from bacterial chemotaxis. *J. Biol. Chem.* **277**, 9625–9628 (2002).
- Spiro, P. A., Parkinson, J. S. & Othmer, H. G. A model of excitation and adaptation in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA* **94**, 7263–7268 (1997).
- Barkai, N. & Leibler, S. Robustness in simple biochemical networks. *Nature* **387**, 913–917 (1997).
- Alon, U. *et al.* Response regulator output in bacterial chemotaxis. *EMBO J.* **17**, 4238–4248 (1998).
- Alon, U., Surette, M. G., Barkai, N. & Leibler, S. Robustness in bacterial chemotaxis. *Nature* **397**, 168–171 (1999).
- Morton-Firth, C. J., Shimizu, T. S. & Bray, D. A free-energy-based stochastic simulation of the Tar receptor complex. *J. Mol. Biol.* **286**, 1059–1074 (1999).
- Morton-Firth, C. J. & Bray, D. Predicting temporal fluctuations in an intracellular signalling pathway. *J. Theor. Biol.* **192**, 117–128 (1998).
- Katz, B. & Miledi, R. Statistical nature of acetylcholine potential and its molecular components. *J. Physiol. (Lond.)* **224**, 665–699 (1972).
- Samuel, A. D. & Berg, H. C. Fluctuation analysis of rotational speeds of the bacterial flagellar motor. *Proc. Natl Acad. Sci. USA* **92**, 3502–3506 (1995).
- Rao, C. V., Wolf, D. M. & Arkin, A. P. Control, exploitation and tolerance of intracellular noise. *Nature* **420**, 231–237 (2002).
- Larsen, S. H., Reader, R. W., Kort, E. N., Tso, W. W. & Adler, J. Change in direction of flagellar rotation is the basis of the chemotactic response in *Escherichia coli*. *Nature* **249**, 74–77 (1974).
- Cluzel, P., Surette, M. & Leibler, S. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science* **287**, 1652–1655 (2000).
- Block, S. M., Segall, J. E. & Berg, H. C. Adaptation kinetics in bacterial chemotaxis. *J. Bacteriol.* **154**, 312–323 (1983).
- Ishihara, A., Segall, J. E., Block, S. M. & Berg, H. C. Coordination of flagella on filamentous cells of *Escherichia coli*. *J. Bacteriol.* **155**, 228–237 (1983).
- Khan, S. & Macnab, R. M. The steady-state counterclockwise/clockwise ratio of bacterial flagellar motors is regulated by protonmotive force. *J. Mol. Biol.* **138**, 563–597 (1980).
- Fahner, K. A. *Studies of Bacterial Flagellar Motors and Filaments*. PhD thesis, Harvard Univ. (1995).
- Yi, T. M., Huang, Y., Simon, M. I. & Doyle, J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA* **97**, 4649–4653 (2000).
- Asakura, S. & Honda, H. Two-state model for bacterial chemoreceptor proteins. The role of multiple methylation. *J. Mol. Biol.* **176**, 349–367 (1984).
- Goldbeter, A. & Koshland, D. E. Jr An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl Acad. Sci. USA* **78**, 6840–6844 (1981).
- Levit, M. N. & Stock, J. B. Receptor methylation controls the magnitude of stimulus-response coupling in bacterial chemotaxis. *J. Biol. Chem.* **277**, 36760–36765 (2002).
- Detwiler, P. B., Ramanathan, S., Sengupta, A. & Shraiman, B. I. Engineering aspects of enzymatic signal transduction: photoreceptors in the retina. *Biophys. J.* **79**, 2801–2817 (2000).
- Elf, J., Paulsson, J., Berg, O. G. & Ehrenberg, M. Near-critical phenomena in intracellular metabolite pools. *Biophys. J.* **84**, 154–170 (2003).
- Viswanathan, G. M. *et al.* Optimizing the success of random searches. *Nature* **401**, 911–914 (1999).
- Spudich, J. L. & Koshland, D. E. Jr Non-genetic individuality: chance in the single cell. *Nature* **262**, 467–471 (1976).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements P.C. is indebted to S. Leibler in whose laboratory at Princeton University this work was begun. The authors thank J. Doyle for pointing out the existence of the power law in the CCW intervals distributions, H. Park for technical help with Labview software, U. Alon for sharing the expression vector pUA4, J. S. Parkinson for the RP8610 strain and the pRR27 vector, and F. Cattaneo for the use of computers. The authors are thankful to R. Albert, C. Guet, T. Griggs, C. Macal, M. North and R. Rosner for discussions and comments on the manuscript. This work was supported partially by the MRSEC programme of the NSF and the Cancer Research Foundation. T.E. acknowledges support from a joint research funding from the US Department of Energy. T.S.S. acknowledges support from NEDO, Ministry of Economy, Trade and Industry of Japan.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.C. (cluzel@uchicago.edu).

The myosin motor in muscle generates a smaller and slower working stroke at higher load

Massimo Reconditi¹, Marco Linari¹, Leonardo Lucii¹, Alex Stewart², Yin-Biao Sun³, Peter Boesecke⁴, Theyencheri Narayanan⁴, Robert F. Fischetti⁵, Tom Irving⁵, Gabriella Piazzesi¹, Malcolm Irving³ & Vincenzo Lombardi¹

¹Laboratorio di Fisiologia, DBAG, Università di Firenze, I-50134 Firenze, and OGG, Istituto Nazionale di Fisica della Materia, Italy

²Rosenstiel Center, Brandeis University, Waltham, Massachusetts 02545, USA

³Randall Division of Cell and Molecular Biophysics, King's College London, London SE1 1UL, UK

⁴European Synchrotron Radiation Facility, F-38043 Grenoble Cedex, France

⁵BioCAT Advanced Photon Source, Argonne, Illinois 60439, USA

Muscle contraction is driven by the motor protein myosin II, which binds transiently to an actin filament, generates a unitary filament displacement or 'working stroke', then detaches and repeats the cycle. The stroke size has been measured previously using isolated myosin II molecules at low load, with rather variable results^{1–4}, but not at the higher loads that the motor works against during muscle contraction. Here we used a novel X-ray-interference technique^{5,6} to measure the working stroke of myosin II at constant load⁷ in an intact muscle cell, preserving the native structure and function of the motor. We show that the stroke is smaller and slower at higher load. The stroke size at low load is likely to be set by a structural limit^{8,9}; at higher loads, the motor detaches from actin before reaching this limit. The load dependence of the myosin II stroke is the primary molecular determinant of the mechanical performance and efficiency of skeletal muscle.

Muscle cells from skeletal and cardiac muscle are composed of many identical functional units called sarcomeres, each of which contains overlapping myosin and actin filaments (Fig. 1a). Muscle shortening is generated by the relative sliding of the two types of filament, driven by the working stroke in the myosin-head domains (Fig. 1b, c). The myosin filament (blue in Fig. 1) is symmetrical about its midpoint and contains two regular arrays of myosin heads (red in Fig. 1). When a muscle fibre is illuminated by a parallel beam of X-rays, the 14.5 nm spacing of the heads in each array produces a strong X-ray reflection called the M3. Interference between the diffracted X-rays from the two head arrays in each filament produces a finely spaced modulation of this reflection, from which the interference distance (ID; Fig. 1) between the two arrays can be measured with a precision of a few angstroms^{5,6}. In principle, the myosin II stroke size in the intact muscle fibre can be determined

directly from the change in ID. In practice, the measurement may be compromised by the compliance of the actin and myosin filaments^{10,11}, which produces additional motions of the myosin heads whenever the force in the filaments changes⁶. Here we have eliminated the effects of filament compliance by feedback control of the force^{7,12}, so that the working stroke is determined directly.

Muscle fibres were activated at constant length to produce an isometric force, T_0 . Then the load was suddenly reduced and clamped at a lower value (Fig. 1e). During the 150 μ s load step (phase 1), the fibre shortens because of the compliance of the myosin heads and the actin and myosin filaments^{6,7,13}. Rapid shortening continues for a few milliseconds after the load step, and this phase 2 shortening is thought to be due to the working stroke⁷. This is followed by a period of much slower shortening (phase 3), before a steady shortening velocity is attained (phase 4)¹².

During isometric contraction (at force T_0), the M3 X-ray reflection (Fig. 1f, L_0 , brown) is split into two main peaks by interference between the myosin-head arrays. The relative intensity R_{M3} of the higher angle/lower angle peak (right peak/left peak in Fig. 1f) was about 0.75. At the start of phase 2 of the velocity transient (L_{2s} , orange), R_{M3} had decreased to about 0.45, and by the end of phase 2 (L_{2e} , pink) it was about 0.2. Thus, myosin heads move towards the midpoint of the filament during both phase 1 and phase 2 (Fig. 1a–c). The elastic strain in the filaments is constant during phase 2, so the decrease in R_{M3} in this period provides a direct measure of the working stroke in the actin-attached myosin heads.

The recovery of R_{M3} during phase 3 (Fig. 1f, L_3 , cyan) signals detachment of myosin heads from actin at the end of their working stroke and their return to axial positions further from the midpoint of the myosin filament (Fig. 1d).

We used this protocol to measure the working stroke at three different loads: 0.25, 0.50 and 0.75 times the isometric force T_0 . At higher loads, both the amplitude and speed of fibre shortening in phase 2 are reduced⁷ (Fig. 2a). X-ray interference shows that these functional properties of the muscle fibre are due to the load dependence of the size and speed of the working stroke in the myosin heads (Fig. 2b). Both the decrease in R_{M3} during phase 2 (open triangles) and its recovery during phase 3 (filled circles) are slower at higher load. The time constant of the decrease in R_{M3} during phase 2 was 0.4 ms at 0.25 T_0 , and 2.0 ms at 0.75 T_0 .

The relationship between R_{M3} and the extent of filament sliding in phases 1 and 2 (Fig. 2c, triangles) was, to a first approximation, independent of the load. This shows that myosin heads remain attached to actin during phases 1 and 2 (Fig. 1a–c); the axial motions of the myosin heads are solely determined by the extent of filament sliding. R_{M3} recovers only during phase 3 (Fig. 2b, c, filled circles), when myosin heads detach at the end of the working stroke. The stroke size can therefore be determined from the extent of filament sliding before R_{M3} recovers (Fig. 2c). Since this recovery occurs over a range of filament sliding, the average stroke size was estimated from the mean filament sliding during R_{M3} recovery, which was about 6.5, 9 and 12 nm at 0.75, 0.50 and 0.25 T_0 , respectively.

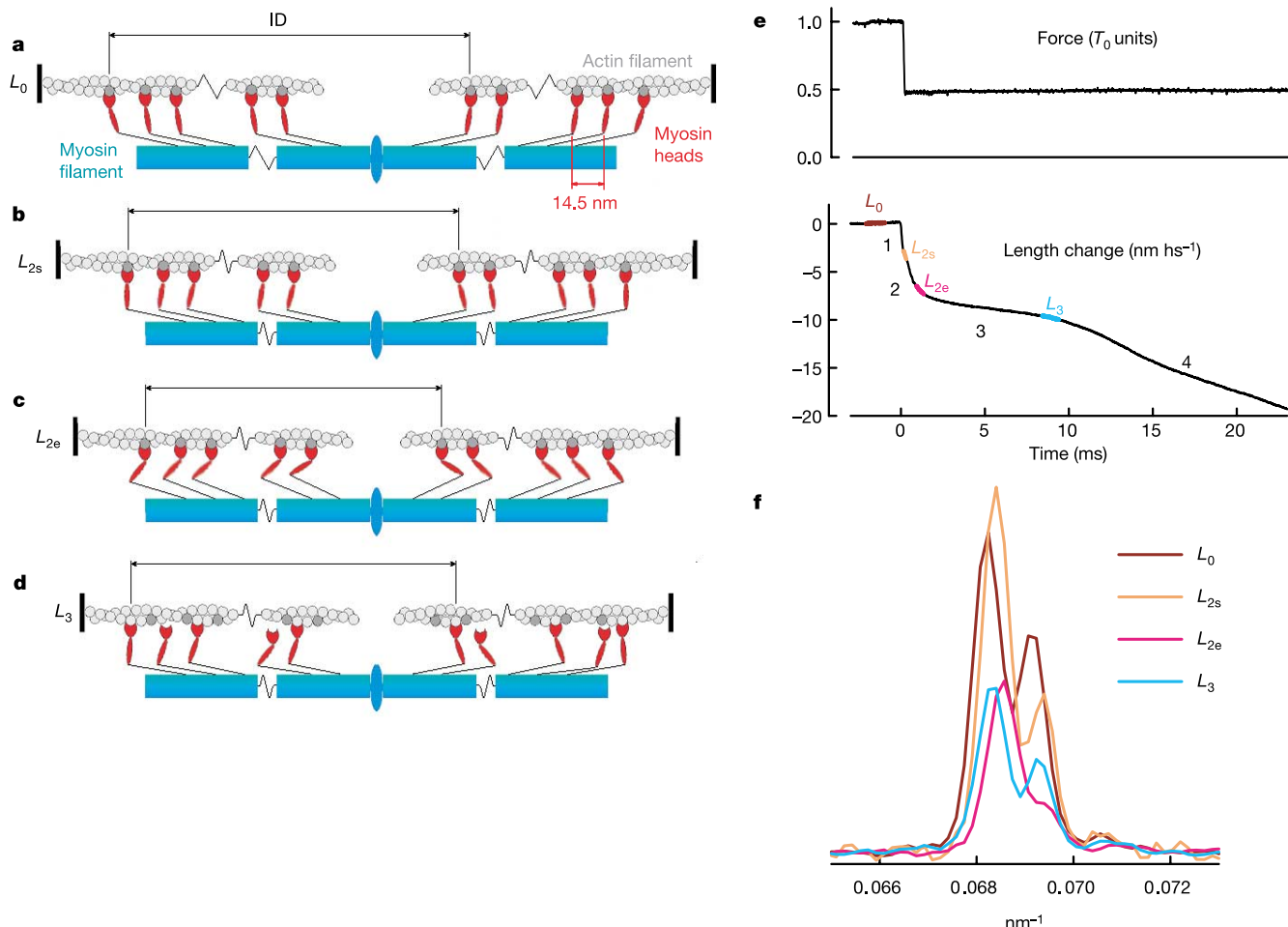


Figure 1 Myosin-head motions following a load step. **a–d**, Organization of myosin and actin filaments and myosin heads in the muscle sarcomere. **e**, Load normalized by isometric force T_0 and length change in nm per half-sarcomere (hs). **f**, Axial X-ray intensity distribution in the region of the M3 reflection recorded at the BioCAT beamline,

normalized by X-ray exposure time. Colours denote X-ray exposure periods shown in **e**: brown, L_0 ; orange, L_{2s} ; pink, L_{2e} ; cyan, L_3 ; total exposure times, 25, 5.5, 21 and 25 ms for the unattenuated beam, respectively. Fibre cross-sectional area, 38,300 μm^2 ; T_0 , 230 kN m⁻²; sarcomere length, 2.16 μm .

The total intensity of the M3 reflection (I_{M3} , Fig. 2d) depends on the axial mass projection of the myosin heads, and therefore on their orientation^{14,15}. The relationship between I_{M3} and filament sliding during phases 1 and 2 was also independent of load. Its shape can be explained by tilting of the actin-attached myosin heads as the filaments slide, with the narrowest mass distribution after about 2 nm of sliding from the isometric conformation^{14,16,17}. I_{M3} recovered slightly during phase 3 (filled circles), as expected from detachment of myosin heads, followed by some reattachment to actin monomers further from the midpoint of the myosin filament (Fig. 1d).

The changes in both R_{M3} and I_{M3} during phases 1 and 2 were quantitatively reproduced by a structural model⁶ in which only half the myosin heads that contribute to the M3 reflection are attached to actin and tilt during filament sliding (see Supplementary Fig. 1). However, the intensity and interference fine structure of the M6 reflection (with a spacing corresponding to the second order of the M3) were almost constant during the velocity transient (see Supplementary Fig. 2), and this observation is not consistent with the published structural model⁶. The analysis of changes in the intensity and interference fine structure of the M3 and M6 reflections during velocity transients presented in Supplementary Information shows that the previous model should be modified by introducing disorder of the myosin heads. The disorder of the force-generating heads corresponds to a uniform angular distribution of their light-chain domains by at least $\pm 17^\circ$. These myosin heads would then make very little contribution to the M6 reflection; this reflection seems to be mainly due to another structural component of the myosin filament^{4,5,14,18}.

The change in the spacing of the M6 reflection (S_{M6}) in a load-step experiment gives a precise measure of the compliance of the myosin filament. S_{M6} in each phase of the velocity transient

depended linearly on the load, and the myosin-filament compliance calculated from these data was $0.26 \pm 0.01\%/T_0$. This is larger than previous estimates^{6,19} based on the changes in the spacing of the M3 reflection (S_{M3}) during phase 1. S_{M3} decreased during shortening at constant load, and in general does not act as the result of a simple compliance⁶, probably because of its origin in the force-generating heads that crosslink the myosin and actin filaments. Previous estimates of myosin-filament compliance based on S_{M3} are therefore likely to be unreliable. The actin-filament compliance is also $0.26\%/T_0$ in the conditions of the present experiments^{19,20}. These values of actin- and myosin-filament compliance, combined with the 5.1 nm/ T_0 compliance of each set of overlapping myosin and actin filaments in the muscle half-sarcomere¹⁹, allow the compliance of the myosin head to be calculated as 1.40 nm/ T_0 . The low compliance of the heads compared with their stroke size leads to important constraints on the motor mechanism^{13,21}.

We have shown that, starting from isometric contraction (force T_0), myosin heads stay attached to actin filaments during filament sliding of 6.5, 9 and 12 nm after load steps to 0.75, 0.50 and 0.25 T_0 , respectively. These distances include the elastic shortening of the half-sarcomere by 5.1 nm/ T_0 during the load step. After subtracting this elastic component, the filament sliding during phase 2 at the three loads was 5.2, 6.4 and 8.1 nm. However, these values exclude the part of the working stroke that may already have been expended in stretching the half-sarcomere compliance by 5.1 nm during the initial generation of the isometric force T_0 before the load step was applied. Thus, the total stroke generated by the myosin II motor *in situ* in each interaction with actin at low load is greater than 8 nm and could be as large as 13 nm. This range centres on the stroke size expected from the difference between the crystallographic conformations of the myosin head with different bound nucleotides^{8,9}, but is more than double that observed in some recent *in vitro* studies of

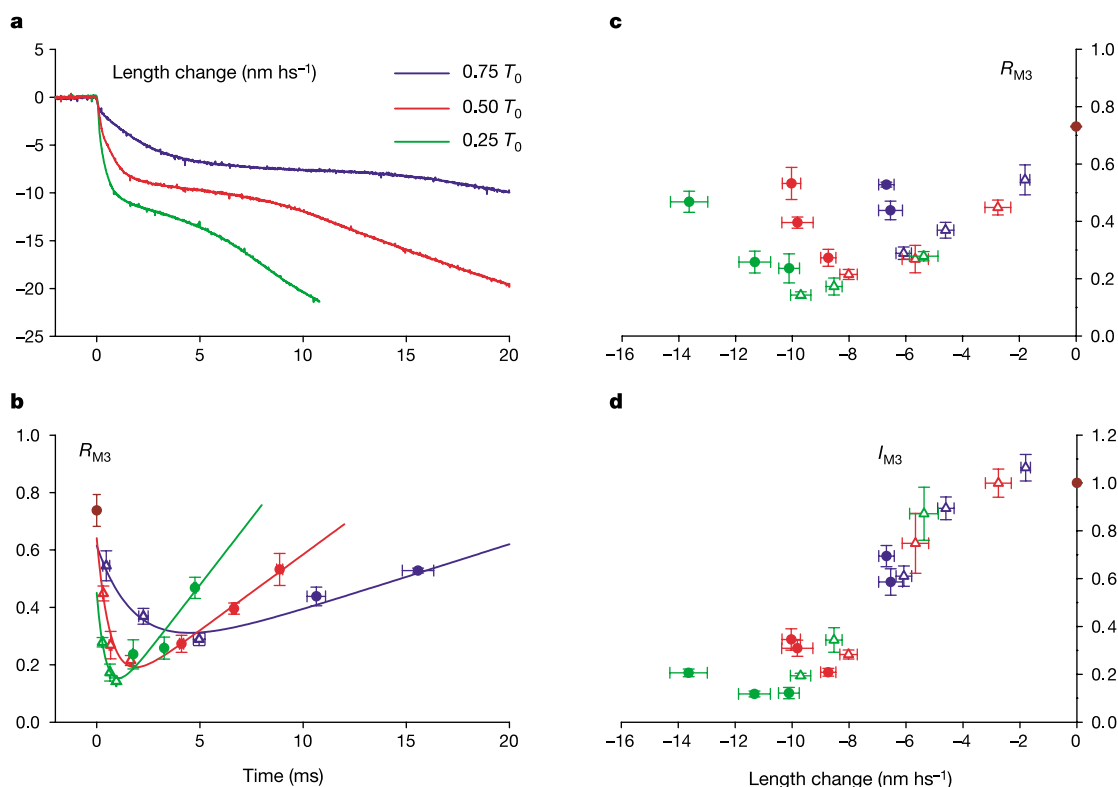


Figure 2 Interference fine structure and intensity of the M3 X-ray reflection. **a**, Fibre length change following load steps to 0.25 T_0 (green), 0.50 T_0 (red) and 0.75 T_0 (blue). **b**, Ratio R_{M3} of the intensities of higher and lower angle peaks of the M3 reflection (right and left peaks, respectively, in Fig. 1f) versus time, as in **a**. **c**, R_{M3} versus extent of

filament sliding. **d**, Total M3 intensity (I_{M3}) normalized to that at L_0 (brown). Data in **b–d** show mean $\pm$ s.e.m. for 4–12 fibres in each class except for L_0 (25 fibres). Open triangles, phase 2; filled circles, phase 3.

isolated myosin II molecules at low load^{2,3}. The latter comparison suggests that those *in vitro* assays underestimate the performance of the motor under physiological conditions, presumably because they do not preserve the native structure and organization of the filaments.

The myosin II motor in skeletal muscle has evolved to generate fast, extensive shortening at low load and efficient work output or tension maintenance at high load. Our results reveal the molecular basis of these basic muscle functions. The stroke size at low load is likely to be determined by a structural limit, corresponding to the distinct myosin-head conformations observed by X-ray crystallography^{8,9}. The stroke size is smaller at high load, and the motor then detaches from actin before reaching the structural limit (Fig. 2c). Both the working stroke and the subsequent detachment are slower at higher load (Fig. 2b). These results provide a molecular explanation for the slower rate of ATP utilization by muscle at higher load^{22,23}. The motor still performs more mechanical work at the higher load, despite the smaller stroke. For example, the work output during phase 2 at 0.75 T_0 is $0.75 \times 5.2 = 3.9 T_0$ nm, which is about twice that (2.0 T_0 nm) at 0.25 T_0 . Because the load is constant in each case, this work cannot be due to the discharge of elastic energy that had been stored previously during isometric contraction. It must be due to active energy transduction in the motor during filament sliding. Thus, our results also reveal the molecular basis of the higher efficiency of energy transduction observed during steady muscle shortening at higher loads^{22,23}. □

Methods

Muscle fibres and mechanical measurements

Frogs (*Rana temporaria*) were cooled to 2–4 °C and killed by decapitation followed by destruction of the spinal cord. Single fibres from the tibialis anterior muscle were mounted between a capacitance force transducer and a loudspeaker-coil motor in Ringer's solution at 4 °C at sarcomere length 2.1 μ m, as described previously²⁴. Sarcomere length in a 1–2 mm fibre segment was measured continuously with a striation follower. Mica windows, 600 μ m apart on either side of the fibre, carried the stimulating electrodes. Fibres were stimulated at 18–25 Hz for up to 2.3 s at 4 min intervals. Two hundred and eighty milliseconds after the first stimulus, the motor control was switched from fixed-end to force-clamp mode, and 20 ms later a force step was imposed in 150 μ s. The parameters of the feedback control were adjusted to optimize the step response for each fibre and load⁷. Data are presented from 25 fibres with cross-sectional areas of $26,000 \pm 7,000 \mu\text{m}^2$ (mean $\pm$ s.d.).

X-ray data collection

Fibres were mounted vertically at the BioCAT beamline 18-ID at the Advanced Photon Source (APS)²⁵ and beamline ID2 of the European Synchrotron Radiation Facility (ESRF)²⁶. Both beamlines provide an X-ray flux of up to about 2×10^{13} photons s^{-1} with a wavelength of approximately 0.1 nm. X-ray exposures were controlled with 10 μ s precision by two electromagnetic shutters in series. The fibre was moved vertically by 100–200 μ m between contractions to spread the effects of radiation damage, which became apparent as a sudden failure of fibre activation or relaxation. Data were typically collected from 20–50 tetani before this occurred.

At the APS, the X-ray beam was focused to full-width half-maximum (FWHM) 50–80 μ m (vertical, V) $\times$ 150 μ m (horizontal, H) on a cooled CCD detector with 7,000 (V) $\times$ 4,000 (H) pixels, 3 m from the fibre²⁷. Pixels were binned by factors of 4 (V) and 32 (H) before read-out after each tetanus. The vertical point-spread function had FWHM 65 μ m. At the ESRF, the X-ray beam had FWHM approximately 100 μ m (V) $\times$ 600 μ m (H) on A3-size storage PhosphorImage plates (Molecular Dynamics) 10 m from the fibre^{5,6,26}. Data were typically accumulated from ten tetani before off-line scanning with a Molecular Dynamics STORM 840 scanner using 100 μ m pixels. The point-spread function of the X-ray camera/scanner combination had vertical FWHM 320 μ m.

Experimental protocol

X-ray data were accumulated from a series of load-step cycles in each tetanus, and in some cases (ESRF experiments) from several tetani in each fibre. For the phase 2 experiments, with the shortest X-ray exposures, 40 shortening/stretch cycles of 50 ms were imposed in each tetanus¹⁷. The duration of each X-ray exposure was 160–800 μ s, depending on the load and time after the load step, and the restretch was about 2, 3 and 6 ms after the load steps to 0.25, 0.50 and 0.75 T_0 , respectively. In the phase 3 experiments, we used 10 or 20 cycles, the restretch was about 5, 10 and 20 ms after load steps to 0.25, 0.50 and 0.75 T_0 , and cycle times were increased to 60, 65 and 80 ms, respectively, to achieve repeatable mechanical responses. X-ray data were also recorded from 1 ms windows before the load step in each cycle, and are shown as the isometric data (L_0 in Fig. 1; points for zero filament sliding or time in Fig. 2). We also recorded diffraction patterns from tetani without imposed length steps, which were similar to those for L_0 .

X-ray data analysis

X-ray diffraction patterns were analysed with HV (A. Stewart), Fit2D (A. Hammersley)

and Peakfit (SPSS Science). Images were centred and aligned using the M3 reflections. The distribution of diffracted intensity along the meridional axis of the X-ray pattern (parallel to the fibre axis) was calculated by integrating from 0.013 nm^{-1} on either side of the axis. The background intensity distribution was fitted using a smooth convex hull algorithm and subtracted. The fine structure of each reflection was analysed by multi-gaussian fitting to obtain the intensity and the peak position of each component. The total intensity of the reflection was determined as the sum of those of the component peaks. The reflection spacing was calculated as the weighted mean of those of the component peaks. Spacings were calibrated using the spacing of the M3 reflection in the resting fibre²⁸, 14.34 nm, or that in an isometric tetanus⁵, 14.573 nm.

Received 31 October 2003; accepted 26 January 2004; doi:10.1038/nature02380.

- Finer, J. T., Simmons, R. M. & Spudis, J. A. Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature* **368**, 113–119 (1994).
- Molloy, J. E. *et al.* Movement and force produced by a single myosin head. *Nature* **378**, 209–212 (1995).
- Metha, A. D. *et al.* Single molecule biomechanics with optical methods. *Science* **283**, 1689–1695 (1999).
- Kitamura, K., Tokunaga, M., Iwane, A. H. & Yanagida, T. A single myosin head moves along an actin filament with regular steps of 5.3 nanometres. *Nature* **397**, 129–134 (1999).
- Linari, M. *et al.* Interference fine structure and sarcomere length dependence of the axial X-ray pattern from active single muscle fibres. *Proc. Natl Acad. Sci. USA* **97**, 7226–7231 (2000).
- Piazzesi, G. *et al.* Mechanism of force generation by myosin heads in skeletal muscle. *Nature* **415**, 659–662 (2002).
- Piazzesi, G., Lucii, L. & Lombardi, V. The size and the speed of the working stroke of muscle myosin and its dependence on the force. *J. Physiol. (Lond.)* **545**, 145–151 (2002).
- Rayment, I. *et al.* Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261**, 50–58 (1993).
- Dominguez, R., Freyzer, Y., Trybus, K. M. & Cohen, C. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualisation of the pre-power stroke state. *Cell* **94**, 559–571 (1998).
- Huxley, H. E., Stewart, A., Sosa, H. & Irving, T. X-ray diffraction measurements of the extensibility of actin and myosin filaments in contracting muscle. *Biophys. J.* **67**, 2411–2421 (1994).
- Wakabayashi, K. *et al.* X-ray diffraction evidence for the extensibility of actin and myosin filaments during muscle contraction. *Biophys. J.* **67**, 2422–2435 (1994).
- Podolsky, R. J. Kinetics of muscular contraction: the approach to the steady state. *Nature* **188**, 666–668 (1960).
- Huxley, A. F. & Simmons, R. M. Proposed mechanism of force generation in striated muscle. *Nature* **233**, 533–538 (1971).
- Huxley, H. E. *et al.* Changes in the X-ray reflections from contracting muscle during rapid mechanical transients and their structural implications. *J. Mol. Biol.* **169**, 469–506 (1983).
- Irving, M., Lombardi, V., Piazzesi, G. & Ferenczi, M. A. Myosin head movements are synchronous with the elementary force-generating process in muscle. *Nature* **357**, 156–158 (1992).
- Piazzesi, G. *et al.* Changes in the X-ray diffraction pattern from single intact muscle fibres produced by rapid shortening and stretch. *Biophys. J.* **68**, 92s–98s (1995).
- Irving, M. *et al.* Confirmation of the myosin motor during force generation in skeletal muscle. *Nature Struct. Biol.* **7**, 482–485 (2000).
- Huxley, H. E., Reconditi, M., Stewart, A. & Irving, T. What the higher order meridional reflections tell us. *Biophys. J.* **84**, 139a (2003).
- Dobbie, I. *et al.* Elastic bending and active tilting of myosin heads during muscle contraction. *Nature* **396**, 383–387 (1998).
- Linari, M. *et al.* The stiffness of skeletal muscle in isometric contraction and rigor: the fraction of myosin heads bound to actin. *Biophys. J.* **74**, 2459–2473 (1998).
- Huxley, A. F. & Tideswell, S. Filament compliance and tension transients in muscle. *J. Muscle Res. Cell Motil.* **17**, 507–511 (1996).
- Kushmerick, M. J. & Davies, R. E. The chemical energetics of muscle contraction. *Proc. R. Soc. Lond. B* **174**, 315–353 (1969).
- Sun, Y.-B., Hilber, K. & Irving, M. Effect of active shortening on the rate of ATP utilisation by rabbit psoas muscle fibres. *J. Physiol. (Lond.)* **531**, 781–791 (2001).
- Lombardi, V. & Piazzesi, G. The contractile response during steady lengthening of stimulated frog muscle fibres. *J. Physiol. (Lond.)* **431**, 141–171 (1990).
- Irving, T. C., Fischetti, R., Rosenbaum, G. & Bunker, G. B. Fibre diffraction using the BioCAT undulator beamline at the Advanced Photon Source. *Nucl. Instrum. Methods Phys. Res. A* **448**, 250–254 (2000).
- Boescke, P., Diat, O. & Rasmussen, B. High-brilliance beamline at the European Synchrotron Radiation Facility. *Rev. Sci. Instrum.* **66**, 1636–1638 (1995).
- Phillips, W. C. *et al.* High-sensitivity CCD-based X-ray detector. *J. Synchrotron Radiat.* **9**, 36–43 (2002).
- Haselgrove, J. C. X-ray evidence for conformational changes in the myosin filaments of vertebrate striated muscle. *J. Mol. Biol.* **92**, 113–143 (1975).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements This work was supported by Ministero dell'Istruzione, dell'Università e della Ricerca, Telethon-945 (Italy), the National Institutes of Health (NIH, USA), the Medical Research Council (UK), the European Molecular Biology Laboratory, the European Union and the European Synchrotron Radiation Facility. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science. BioCAT is an NIH-supported research centre. We thank A. Aiazzi, M. Dolfi and J. Gorini for mechanical and electronics support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to V.L. (vincenzo.lombardi@unifi.it).

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs Sales Director:

Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Scandinavia/ Spain/ Portugal:

Evelina Rubio Håkansson (4973)

Natureevents: Silje Opstrup (4994)

France/ Switzerland:

Amelie Pequignot (4974)

Production Manager:

Billie Franklin
To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Stefan Hales

European Satellite Office

Germany/ Austria/ Italy/

The Netherlands/ Belgium:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager:

Peter Bless

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Haraikatamachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami
e-mail: rasami@naturejpn.com

naturejobs

Exploring careers

While chatting with graduate students from England, Switzerland, Germany and the Netherlands last month, I repeatedly heard one complaint. Surprisingly it wasn't about money. Rather, the students said that they heard little about career paths outside academia, and that the information they did receive seemed murky and ill-defined.

This issue is especially pronounced in Europe, as there is no centralized clearing house for career information there. Several student and postdoc organizations in individual countries are making inroads, but the results, so far, have been fragmented. The European Commission has also been improving its offerings, with efforts such as the 'mobility portal' (see *Nature* 427, 868-869; 2004), but that website tends to focus on moving from country to country rather than general career information for scientists who want to work outside academia.

The US National Institutes of Health may have inadvertently provided at least a partial solution. Last month, it launched a website called LifeWorks (<http://science.education.nih.gov/LifeWorks>), which profiles more than 100 careers in the health and medical sciences.

Although LifeWorks is designed with middle- and high-school students in the United States in mind, it may actually prove to be a boon to young scientists in Europe and beyond. The ability to browse careers by salary, education and job title is relevant to job-seekers of any age. And reading about others' career successes in non-traditional scientific fields transcends boundaries.

Although we hope that *Naturejobs* provides some of the same advice and services, we recognize that the more information on careers that is out there, the better.

Paul Smaglik
Naturejobs editor



Contents

CAREERS & RECRUITMENT

Seeking out skills
for bioinformatics p584

CAREER VIEW

Nuts & Bolts
Interview preparation
Graduate Journal
Undertaking undergraduates
Movers
Vishva Dixit p586

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

Changing of the guard

Opportunities in bioinformatics once abounded for the self-taught and industrially minded, but employers are now turning towards the formally trained and academics. Myrna Watanabe reports.

Francis Ouellette, director of the University of British Columbia (UBC) Bioinformatics Centre, calls his generation of computational biologists “the old guard”. Ouellette quit his doctoral programme in developmental biology after dabbling in informatics turned him into “the lab computer expert”. After stints managing the yeast-chromosome project and serving as GenBank coordinator for the US National Center for Biotechnology Information in Bethesda, Maryland, Ouellette moved in 1998 to UBC, where he is now a full professor — but not in the discipline in which he originally trained.

Lincoln Stein, an associate professor at Cold Spring Harbor Laboratory, is also a member of Ouellette’s old guard. A pathologist by training, Stein moved from autopsies to bioinformatics while doing computer consulting in medical school. Ironically, when Stein is hiring people for his group, he doesn’t turn to self-trained types such as himself or Ouellette. Instead, he looks for people who have spent a number of years in bioinformatics software development. “Now there is a pool of experienced individuals,” says Stein, “and those are the people that I prefer.”

The bioinformatics field has changed during the years since it first developed. It grew from the need to build gene-sequence databases and write the software to mine them. But the deluge of data from technologies such as automated sequencing, microarrays and mass spectrometry attracted a diverse range of professionals who now manage, manipulate and interpret the new information. Along with people such as Ouellette and Stein, who developed

bioinformatics and continue to define its new directions, the field now includes IT professionals who have little or no biology background, and a generation specially trained in bioinformatics.

The job market has also changed. Before the bubble burst in the 1990s, big drug companies might have employed up to 100 bioinformatics specialists, and plenty more worked in biotechnology start-ups. In Asia, government labs and corporations are still hiring workers in bioinformatics, says Santosh Mishra, executive director of the Bioinformatics Institute in



Another dimension

The completion of the rat genome sequence this month (see page 493) will not only provide bioinformaticians with another data set for their efforts in ‘comparative genomics’, it will also add a new dimension to available data. Because the rat is used as a model organism for physiology, bioinformaticians will be able to discern how genes affect different

tissues and processes. Matching physiological traits to genetic maps of the rat will be an exciting challenge and an area of immense opportunity.

“The bioinformatics is going to require people who can come in and find a new way to display data,” says Howard Jacob, director of the Human and Molecular Genetics Center at the Medical College of

Wisconsin in Milwaukee, and a co-author on the rat genome paper.

Jacob adds that scientists who want to move into systems biology would do well to look at the rat. Learning to work with both physiological and genetic data — and perhaps proteomics data — will help scientists to model how disease processes work in the whole organism, Jacob says. **Paul Smaglik**

Singapore. But in Europe and the United States, more opportunities are to be found in academia. Some US universities are advertising for as many as 20 positions in the field.

Former industry employees are moving over in droves. Computational biologist Gene Myers, for example, the former vice-president of informatics research at Celera Genomics, is now a professor at the University of California, Berkeley.

THE VANGUARD

The universities, for their part, are looking out for members of the new 'vanguard', says Susan Davidson, interim director of the University of Pennsylvania's Center for Bioinformatics — people who are training to develop the next gene-finding tools or who are building software to solve problems that have only recently been

working on his postdoc. "In a year or two he'll be very much in demand because he has covered all areas," Tozeren says.

Although the up-and-coming PhD practitioners are preparing to lead, an even bigger pool of highly skilled technicians is training to help. This group — generally pitched at master's level — may not be so concerned with determining the next big problems in biology. Instead, their role is to help solve them, through a combination of databases, software and hardware, says Joyce Mitchell, director of the training programme for biomedical and health informatics research at the University of Missouri in Columbia. "They're contributing but they're not leading," she says.

Around the world, programmes are being developed to train mid-level bioinformaticians. Roland Eils, of the German Cancer Research Centre in

Heidelberg, notes that for the past three years, the University of Heidelberg has been running a one-year, extended weekend course that gives graduates in biology, computational science or mathematics a certificate in bioinformatics. "These training programmes are suited to the down-to-earth needs of bioinformatics in industry," Eils says.

Ripples are also being felt from the undergraduate level. The University of California, Santa Cruz, has a programme that includes a bioinformatics degree under the tutelage of David Haussler, director

of the university's centre for biomolecular science and engineering.

Perhaps the biggest wild card affecting employment in the bioinformatics pool is whether it will take in one displaced group — downsized information-technology workers. Tozeren says that these people need to pick up additional skills if they want to compete with the specialists, whatever the degree level. But courses exist to help with this conversion. For example, the Greater Philadelphia Bioinformatics Alliance offers two- or three-day intensive workshops to introduce informatics computer programs, for example, or microarray data analysis.

There are also signs that IT professionals without a biology background can play a role. Nicholas Ide, president of Thoughtful Solutions in McLean, Virginia, is one. His company works for the US National Institutes of Health on what he calls "the information domain related to medicine", developing databases on clinical trials for the public to use.

So much information is now available, and needing people to handle it, that skilled professionals of various backgrounds should still be able to find work if they're adaptable. But in a field changing so quickly, they will all need to keep updating their skills so that they are not eclipsed by the next generation. ■

Myrna Watanabe is a freelance writer based in Patterson, New York.



The old guard: bioinformatics was not part of the master plan for either Francis Ouellette (left) or Lincoln Stein.

identified. This group has the advantage of directly pursuing this role, rather than falling into it.

Geneticist Aaron Chang, from the University of Washington Medical School in Seattle, is a member of this vanguard. By the end of his postdoc, he will be highly trained in both biology and informatics. His postdoctoral programme covers bioinformatics related to genetic data and medical informatics, handling patient data at the bedside.

There are jobs for people with Chang's training, says Janet Thornton, director of the European Bioinformatics Institute in Hinxton, near Cambridge, UK. Her institute, for instance, hires group leaders after they have completed one or two postdocs, she says.

In some areas, you don't even need that. The market for such people is so good, says Sean Eddy of the department of genetics at Washington University in St Louis, Missouri, that some jobs go to people without postdocs. "I know of one person who did a three-year PhD and then moved out into a faculty position," says Eddy.

But a good range of skills can still be valued. Aydin Tozeren, director of bioinformatics at Drexel University in Philadelphia, Pennsylvania, explains that he has a student who worked as an engineer for 15 years, who then went to medical school, and is now

GRADUATE JOURNAL

Undertaking undergraduates

One spring, when I was an earnest and energetic young master's student, an undergraduate asked if she could work on a summer project with me. Thoughts of generating twice the data per unit time and having someone else brew the lab coffee fleetingly crossed my mind. Why not?

The idea of an immediate increase in data evaporated when I found out how time-consuming it is to bring an undergraduate up to speed. I wanted her experience in research to be good, so I organized a four-month project for her, briefed her on the wider scope of her work and checked her results. It was harder than I thought to make a series of repeated experiments seem important — how do you motivate a student when research is often quite dry? My own work slowed down.

In the end, we generated a fairly complete report. The student was happy, although I had accomplished less that summer than I had wanted. But I learned a lot about organizing someone else's project, and it became somewhat addictive to show students how rewarding research can be.

I have since supervised other summer projects. But not this year — I need to focus on finishing my own work. It's time for another fresh graduate student to figure out their supervising style. ■

Sidney Omelon is a PhD student in bone biomaterials at the Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada.

NUTS & BOLTS

Interview preparation

It's the moment you've been waiting for. You have managed to get an interview for a coveted position. The next step is to show up and dazzle them with your qualifications and charm, right? Not quite so fast.

An interview is as much about your potential employer as it is about you. So you need to do some research. Granted, this isn't a typical scientific assignment, but finding out as much as you can about your possible workplace can only help you in the interview.

The first step is obvious — hit the organization's website for an overview. This applies for posts in academia as well as industry. Most academic departments have their own website, as do many principal investigators. It's good to familiarize yourself with the work of potential colleagues before you meet them — whatever the sector.



With Deb Koen
Careers consultant

After you've seen what the organization has to say for itself, move on to find out how others view it. You can get corporate information from directories such as Standard & Poor's Register of Corporations, Hoover's Online and D&B Million Dollar Database. And Vault.com goes beyond financial data to offer company snapshots.

For an outside perspective of academia, publications such as *The Chronicle of Higher Education* provide a good overview of the US system. Academic society

publications and scientific journals also offer timely coverage of issues and organizations.

But the best source of knowledge comes from people who have worked for the organization. Try talking to someone who holds a similar position to the one you are seeking, or contact former employees who held the position in the past.

Continue researching during the interview by talking to your prospective co-workers. Find out what they spend their time on, as well as what they relish and abhor about their jobs. Soak in all you can about the position and the culture to imagine yourself working there.

Such thorough research will impress the interviewer with your knowledge and, more importantly, will allow you to judge your fit with the organization. ■

Deb Koen is vice-president of Career Development Services and a columnist for *The Wall Street Journal's CareerJournal.com*.

MOVERS

Vishva Dixit, vice-president of molecular oncology, Genentech, South San Francisco



Vishva Dixit almost gleefully chose a career path that ran contrary to conventional wisdom. Raised in a strong family tradition of medicine (both of his parents practised in Kenya when he was growing up), he traded his MD for a research position focusing on molecular biology. Rather than doggedly pursue an early interest, he switched to another topic after reading a popular magazine. And after climbing the academic ladder to

tenure, he left for industry. Once there, he thought he'd be content to while away his career doing research, but late last year he joined the ranks of senior management.

Dixit's first change, when he chose basic research over clinical practice, mortified his mother. "She always said: 'Why did you go to medical school if you're going to do this?'," Dixit says.

But he was fascinated by cellular processes, and pursued a line of research in extracellular matrices. "Then one day, out of the blue, I decided to change focus," he says. "I just wasn't making inroads."

The defining moment came when he read an article in *Scientific American* about tumour-necrosis factor. Back then, this protein was noted for its association with inflammation, but Dixit was struck by its ability to kill cells. So he contacted the biotech firm Genentech in South San Francisco, which was doing research on the protein and could produce large

quantities of it, and asked for some samples. "The company was exceedingly generous and sent me a thousand times more than I needed," Dixit says.

Those samples allowed Dixit and his colleagues to tease apart the mechanisms of programmed cell death, or apoptosis, and led to Genentech inviting him to a job interview. At first, Dixit wasn't interested — he planned to see his career out as a tenured professor at the University of Michigan. But he thought he would take advantage of the flight to the Bay Area to visit his brother.

Once at the company's offices, he was impressed by its commitment to basic research, but was still wary. He had what he describes as "the academic's traditional suspicion of industry". So he talked to David Goeddel, one of Genentech's scientific founders, who had left to run Tularik, a Genentech spin-off. He helped to convince Dixit to go against the grain one more time. ■

CV

1999–: Adjunct professor, pharmaceutical-chemistry department, University of California, San Francisco

1997–2003: Director and then senior director of molecular oncology, Genentech, South San Francisco

1986–99: Assistant professor rising to professor in 1995, pathology department, University of Michigan Medical School, Ann Arbor, Michigan

1982–86: Postdoctoral fellow, Department of Biological Chemistry, Washington University School of Medicine, St Louis, Missouri